

1/33

SIGNAL PEPTIDE
HUMAN -19 MQIELSTCFF LCLLRFCFS
PIG MQLELSTCVF LCLLPLGFS
MOUSE MQIALFACFF LSLFNFCSS
* * * * *

FIG. 1A

A1 DOMAIN
HUMAN 1 ATRRYYLGA V ELSWDYMQSD LG-ELPVDAR FPFRVPKSF FNTSVVYKKT
PIG AIRRYYLGA V ELSWDYRQSE LLRELHVDTR FPATAPGALP LGPSVLYKKT
MOUSE AIRRYYLGA V ELSWNYIQSD LLSVLHTDSR FLPRMSTSFP FNTSIMYKKT
* * * * *

50 LFVEFTDHLF NIAKPRPPWM GLLGPTIQAE VYDTVVITLK NMASHPVSLH
VFVEFTDQLF SVARPRPPWM GLLGPTIQAE VYDTVVITLK NMASHPVSLH
VFVEYKDQLF NIAKPRPPWM GLLGPTIWTE VHDTVVITLK NMASHPVSLH
* * * * *

100 AVGVSYWKAS EGA EYDDQTS QREKEDDKVF PGGSHTYVWQ VLKENGPMAS
AVGVSFWKSS EGA EYEDHTS QREKEDDKVL PGKSQTYVWQ VLKENGPTAS
AVGVSYWKAS EGA EYEDQTS QMEKEDDKVF PGESHTYVWQ VLKENGPMAS
* * * * *

150 DPLCLTYSYL SHVDLVKDLN SGLIGALLVC REGSLAKEKT QTLHKFILLF
DPPCLTYSYL SHVDLVKDLN SGLIGALLVC REGSLTRERT QNLHEFVLLF
DPPCLTYSYM SHVDLVKDLN SGLIGALLVC KEGSLSKERT QMLYQFVLLF
* * * * *

200 AVFDEGKSWH SETKNSLMQD RDAASARAWP KMHTVNGYVN RSLPGLIGCH
AVFDEGKSWH SARND SWTRA MDPAPARAQP AMHTVNGYVN RSLPGLIGCH
AVFDEGKSWH SETNDSYTQS MDSASARDWP KMHTVNGYVN RSLPGLIGCH
* * * * *

250 RKSVYWHVIG MGTTP EVHSI FLEGHTFLVR NHRQASLEIS PITFLTAQTL
KKS VYWHVIG MGTSPEVHSI FLEGHTFLVR HHRQASLEIS PLTFLTAQTF
RKSVYWHVIG MGTTP EIHSI FLEGHTFFVR NHRQASLEIS PITFLTAQTL
* * * * *

300 LMDLGQFLLF CHISSHQHDG MEAYVKVDSC APC/IXa NEEAEDYDDD
LMDLGQFLLF CHISSHHHG MEAHVRVESC AEEPQLRRKA DE-EEDYDDN
LIDLGQFLLF CHISSHKHDG MEAYVKVDSC PEESQWQKKNN EEMEDYD
* * * * *

350 LTDSEMDVVR FDDDNSPSFI QIR I Ia/Xa
LYDSMDVVR LDGDDVSPFI QIR
DDL YSEMDMF TLDYDSSPFI QIR
* * * * *

FIG. 1B

2/33

A2 DOMAIN

HUMAN	373	SVAKKHPKTW	VHYIAAEEED	WDYAPLVLP	DDRSYKSQYL	NNGPQIRGRK
PIG		SVAKKHPKTW	VHYISAEED	WDYAPAVPSP	SDRSYKSLYL	NSGPQIRGRK
MOUSE		SVAKKYPKTW	IHYISAEED	WDYAPSVPTS	DNGSYKSQYL	SNGPHRIGRK
		*****	*****	*****	*****	*****

423	YKKVRFMAYT	DETFKTREAI	QHESGILGPL	LYGEVGDPLL	IIFKNQASRP
	YKKARFVAYT	DVTFKTRKAI	PYESGILGPL	LYGEVGDPLL	IIFKNKASRP
	YKKVRFIAYT	DETFKTRETI	QHESGLLGPL	LYGEVGDPLL	IIFKNQASRP
	***	***	***	***	***

A2 INHIBITOR EPITOPE

473	YNIYPHGITD	VRPLYSRRLP	KGVKHLKDFP	ILPGEIFKYK	WTVTVEDGPT
	YNIYPHGITD	VSALHPGRLL	KGWKHLKDMP	ILPGETFKYK	WTVTVEDGPT
	YNIYPHGITD	VSPLHARRLP	RGIKHVKDLP	IHPGEIFKYK	WTVTVEDGPT
	*****	***	***	***	***

F.IXa BINDING

APC

523	KSDPRCLTRY	YSSFVNMERD	LASGLIGPLL	ICYKESVDQR	GNQIMSDKRN
	KSDPRCLTRY	YSSSINLEKD	LASGLIGPLL	ICYKESVDQR	GNQIMSDKRN
	KSDPRCLTRY	YSSFINPERD	LASGLIGPLL	ICYKESVDQR	GNQIMSDKRN
	*****	***	***	***	***

573	VILFSVFDEN	RSWYLTENIQ	RFLPNPAGVQ	LEDPEFQASN	IMHSINGYVF
	VILFSVFDEN	QSWYLAENIQ	RFLPNPDGLQ	QDPEFQASN	IMHSINGYVF
	VILFSIFDEN	QSWYITENMQ	RFLPNAAKTQ	QDPPGFQASN	IMHSINGYVF
	*****	***	***	***	***

623	DSLQLSVCLH	EVAYWYILSI	GAQTDFLSVF	FSGYTFFHKM	VYEDTLTLFP
	DSLQLSVCLH	EVAYWYILSV	GAQTDFLSVF	FSGYTFFHKM	VYEDTLTLFP
	DSLELTVCLH	EVAYWHILSV	GAQTDFLSIF	FSGYTFFHKM	VYEDTLTLFP
	***	***	***	***	***

◆◆

673	FSGETVFMSM	ENPGLWILGC	HNSDFRNRGM	TALLKVSSCD	KNTGDYYEDS
	FSGETVFMSM	ENPGLWVLGC	HNSDLRNRGM	TALLKVYSCD	RDIGDYDNT
	FSGETVFMSM	ENPGLWVLGC	HNSDFRKRGM	TALLKVSSCD	KSTSDYEEI
	*****	*****	***	***	***

IIa/Xa/APC

723	YEDISAYLLS	KNNAIEPR
	YEDIPGFLS	GKNVIEPR
	YEDIPTQLVN	ENNVIDPR
	***	***

FIG. 1C

3/33

B DOMAIN

HUMAN	741	SFSONSRHPS	TROKQFNATT	IPENDIEKTD	PWFAHRTMP	KIQNVSSSDL
PIG		SFAQNSRPPS	ASQKQFQTIT	SPEDDVE-LD	POSGERTQAL	EELSVPSGDG
MOUSE		SFFQNTNHPN	TRKKKFKDST	IPKNDMEKIE	PQFEEIAEML	KVQSVSVSDM
		** ** *	* * *	* * **	*	* *
	791	LMLLRQS-PT	PHGLSLSDLQ	EAKYETFSDD	PSPGAIDSNN	SLSEMTHFRP
		SMLLGQN-PA	PHGSSSSDLQ	EARNEA--DD	YLPGARERNT	APSAAARLRP
		LMLLGQSHPT	PHGLFLSDGQ	EAIYEAIHDD	HSPNAIDSNE	GPSKVTQLRP
		*** *	*** * *	** * **	* * *	* **
	840	QLHHSGDMVF	TPESGLQLRL	NEKLGTTAAT	ELKKLDFKVS	ST-SNNLIS-
		ELHHSARVL	TPEP-----	-----EK	ELKKLDSKMS	SSSDLLKTSP
		ESHHSKIVF	TPQPGQLQRS	NKSLETTIEV	KWKKLGLQVS	SLPSNLMTT-
		*** *	**		*** *	*
	888	TIPSDNLAAGT	DNTSSLGPPS	MPVHYDSOLD	TTLFGKKSSP	LTESGGPLSL
		TIPSDTLAET	ERTHSLGPPH	PQVNFERSQLG	AIVLGKNSSH	FIGAGVPLGS
		TILSDNLKATF	EKTDSSGFPD	MPVHSSSKLS	TTAFGKKAYS	LVGSHVPLNA
		** * * *	* * * *	* * *	**	**
	939	SEENNDKLL	ESGLMNSQES	SWGKNVSSTE	SGRLFKGKRA	HGPALLTKDN
		TEED-----	-----HES	SLGENVSPVE	SDGIFEKERA	HGPASLTKDD
		SEENSDSNIL	DSTLMYSQES	LPRDNILSIE	NDRLREKRF	HGIALLTKDN
		**	**	*	*	** * ****
	989	ALFKVISISLL	KTNKTSNNSA	TNRKTHIDGP	SLLIENSPSV	WQNILESDE
		VLKVNISLV	KTNKARVYLK	TNRKIHIDDA	ALLTENRAS-	-----
		TLFKDNVSLM	KTNKTYNHST	TNEKLHTESE	TSIENSTTDL	QDAILKVNSE
		*** **	****	** * *		
	1039	FKKVTPLIHD	RMLMDKNATA	LRLNHMSNKT	TSSKNMEMVQ	QKKEGPIPPD
		-----	ATFMDKNTTA	SGLNHVSN--	-----	-----
		IQEVTALIHD	GTLLGKNSTY	LRLNHMLNRT	TSTKNKDIFH	RKDEDPIPOD
		* ***	**	*** *		
	1089	AQNPDMSEFFK	MLFLPESARW	IQRTHGKNSL	NSGQGPSPKQ	LVSLGPEKSV
		-----	-----W	IKGPLGKNPL	SSEGRPSPEL	LTSSGSGKSV
		EENTIMPFSK	MLFLSESSNW	FKKTNGNNSL	NSEQEHSPKQ	LVYLMFKKYV
			*	* * *	* **	* *
	1139	EGQNFLSEKN	KVVVGKGEFT	KDVGLKEMVF	PSSRNLFILT	LDNLHENNTH
		KGQSSGQGR	RVAVEEEELS	KG---KEMML	PNSELTFLT	SADVQGNTH
		KNQSFLEKN	KVTVEQDGFT	KNIGLKDMAF	PHNMSIFLT	LSNVHENGRH
		*	* *	*	***	* *
	1189	NQEKKIQEEI	EKKETLIQEN	VVLPOIHTVT	GTKNFMKNLF	LLSTRQNV
		SQGKKSREEM	ERREKLVQEK	VDLPQVYTAT	GTKNFLRNIF	HQSTEPSV
		NQEKNIQEEI	EK-EALIEEK	VVLPOVHEAT	GSKNFLKDIL	ILGTRQNI--
		* *	* * * *	* **	* **	
	1239	SYDGAYAPVL	QDFRSLNDST	NRTKKHTAHF	SK--KGEEN	LEGLGNQTKQ
		FDGGSHAPVP	QDSRSLNDSA	ERAETHIAHF	SAIR--EAP	LEAPGNRT--
		SLYEVHVPVL	QNITSINNST	NTVQIHMEHF	FKRRKDKETN	SEGLVNKTRE
		**	* * * *	* **	*	* *

FIG. 1D-1

FIG. 1D-2

EIGHT CHAIN ACTIVATION PEPTIDE I Ia/Xa
 HUMAN 1649 EITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPR
 PIG
 DISLPTFQPEEDKMDYDDIFSTETKGEDFDIYGEDENQDPR
 MOUSE
 EL--SAFQSEQEATDYDDAITIET-IEDFDIYSEDIKQGPR
 * * * * *
 * * * * *

FIG. 1E

5/33

A3 DOMAIN

				IXa	Xa	
HUMAN	1690	SFQKKTRHYF	IAAVERLWDY	GMSSSPHVLR	NRAQSGSVPO	FKKVVFQEF
PIG		SFQKRTRHYF	IAAVEQLWDY	GMSESPRALR	NRAQNGEVPR	FKKVVFREF
MOUSE		SVQQKTRHYF	IAAVERLWDY	GMSTS-HVLR	NRYQSDNVPQ	FKKVVFQEF
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
	1740	DGSFTQPLYR	GELNEHLGLL	GPYIRAEVED	NIMVTFRNQA	SRPYSFYSSL
		DGSFTQPSYR	GELNKHLLG	GPYIRAEVED	NIMVTFRNQA	SRPYSFYSSL
		DGSFSTQPLYR	GELNEHLGLL	GPYIRAEVED	NIMVTFRNQA	SRPYSFYSSL
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
				FACTOR IXa BINDING		
	1790	ISYEEDOROG	AEPRKNEVKP	NETKTYFWKV	QHHMAPTKDE	FDCKAWAYFS
		ISYPDDQEQG	AEPRHNFVQP	NETRTYFWKV	QHHMAPTEDE	FDCKAWAYFS
		ISYKEDQR-G	EEPRRNFVKP	NETKIYFWKV	QHHMAPTEDE	FDCKAWAYFS
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
	1840	DVDLEKDVHS	GLIGPLLICH	TNTLNPAHGR	QVTVQEFALF	FTIFDETKSW
		DVDLEKDVHS	GLIGPLLICH	ANTLNAAHGR	QVTVQEFALF	FTIFDETKSW
		DVDLERDMHS	GLIGPLLICH	ANTLNPAHGR	QVSVQEFALL	FTIFDETKSW
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
	1890	YFTENMERN	RAPCNQMED	PTFKENYRFH	AINGYIMDTL	PGLVMAQDQR
		YFTENVERN	RAPCHLMED	PTLKENYRFH	AINGYVMDTL	PGLVMAQNQR
		YFTENVKRNC	KTPCNFQMED	PTLKENYRFH	AINGYVMDTL	PGLVMAQDQR
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
	1940	IRWYLLSMGS	NENIHSIHFS	GHVFTVRKKE	EYKMALYNLY	PGVFETVEML
		IRWYLLSMGS	NENIHSIHFS	GHVFSVRKKE	EYKMAVYNLY	PGVFETVEML
		IRWYLLSMGN	NENIQSIHFS	GHVFTVRKKE	EYKMAVYNLY	PGVFETLEMI
		* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
				PROTEIN C BINDING		
	1990	PSKAGIWRVE	CLIGEHLHAG	MSTLFLVYSN		
		PSKVGIWRIE	CLIGEHLQAG	MSTLFLVYSK		
		PSRAGIWRVE	CLIGEHLQAG	MSTLFLVYSK		
		* * * * *	* * * * *	* * * * *		

FIG. 1F

```

2120  TGTLMVFFGN  VDSSGIKHN  FNPPIIARYI  RLHPTHYSIR  STLRMELMGCDLN
      TGTLMVFFGN  VDASGIKHN  FNPPIVARYI  RLHPTHYSIR  STLRMELMGCDLN
      TGTLMVFFGN  VDSSGIKHS  FNPPIIARYI  RLHPTHSSIR  STLRMELMGCDLN
      *****  **  *****  *****  *****  *****  *****

```

FIG. 1G

2323 *BINDING*
VLGCEAODLY
VLGCEAODLY
ILGCEAQQQY
***** *

FIG. 1H

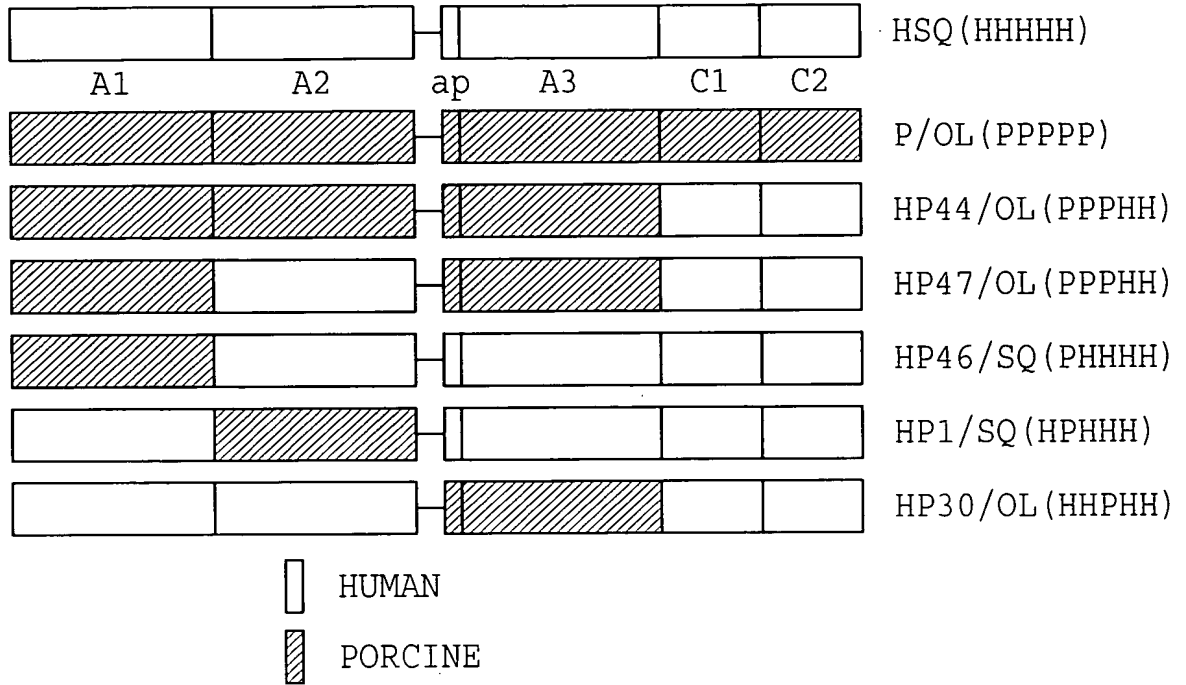


FIG. 2

8/33

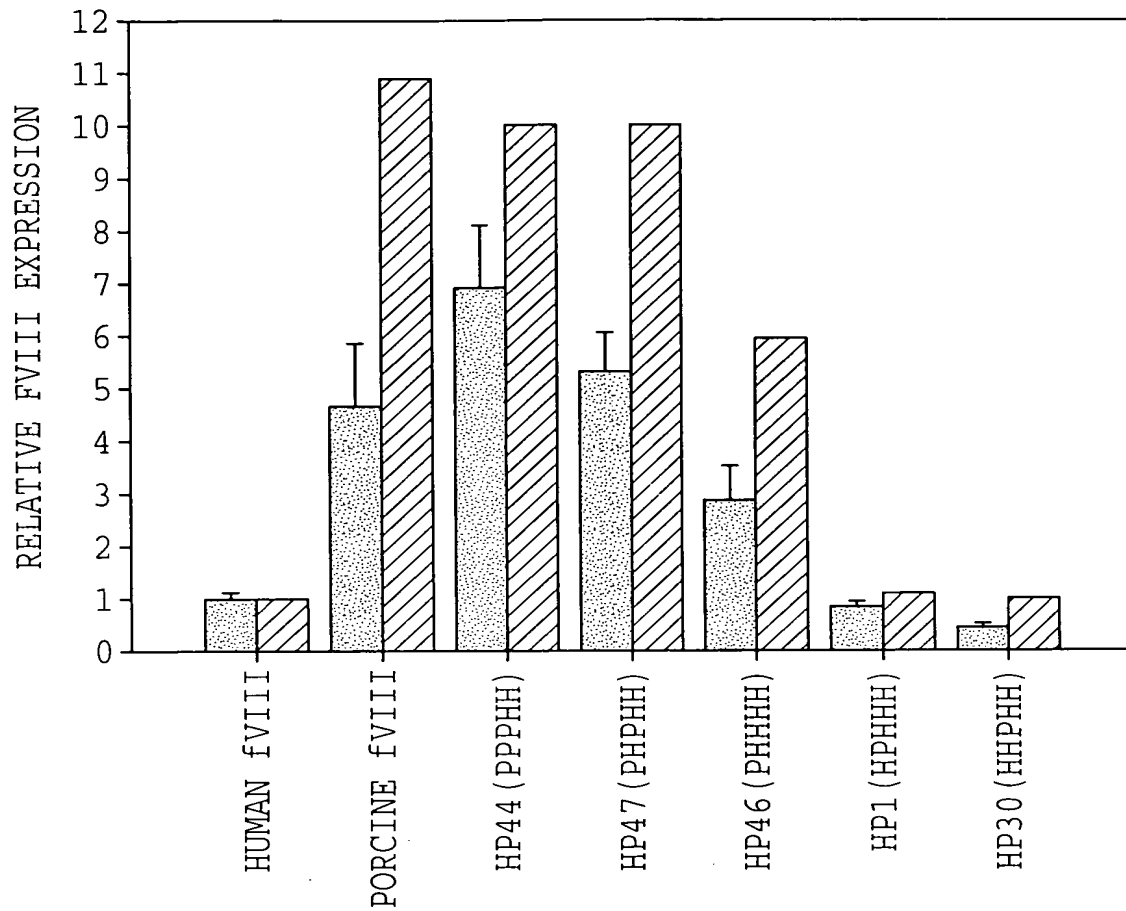


FIG. 3

9/33

AMINO ACID SEQUENCE OF HP44/OL

1	MQLELSTCVF	LCLLPLGFSA	IRRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYVWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLFC	HISSHHHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRDL	GDDVSPFIQI	RSVAKKHPKT
401	WVHYISAEED	DWDYAPAVPS	PSDRSYKSLY	LNSGPQRIGR	KYKKARFVAY
451	TDVTFKTRKA	IPYESGILGP	LLYGEVGDTL	LIIFKNKASR	PYNIYPHGIT
501	DVSALHPGRL	LKGWKHLKDM	PILPGETFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSSINLEK	DLASGLIGPL	LICYKESVDQ	RGNQMMSDKR	NVILFSVFDE
601	NQSWYLAENI	QRFLPNPDGL	QPQDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYILS	VGAQTDFLSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWVLG	CHNSDLNRNG	MTALLKVYSC	DRDIGDYDYN	TYEDI PGFLL
751	SGKNVIEPRS	FAQNSRPPSA	SAPKPPVLR	HORDISLPTF	QPEEDKMDYD
801	DIFSTETKGE	DFDIYGEDEN	QDPRSFOKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALNRQAQNG	EVPRFKKVVF	REFADGSFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVTF	KNQASRPYSF	YSSLISYPDD	QEQAEPHRN	FVQPNETR TY
951	FWKVQHMAP	TEDEFDC KAW	AWFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLKEN
1051	YRFHAINGYV	MDTLPGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFSV
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVG I	WRIECLIGEH	LQAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASGOYGO	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKV DL	LAPMIIHG I K	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNIFNPPII	ARYIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLQGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISSS
1401	QDGHQWTLFF	QNGKVKV FQG	NQDSFTPVVN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY*			

1-19 SIGNAL PEPTIDE
20-391 A1 DOMAIN
392-759 A2 DOMAIN
760-783 OL LINKER
784-1154 ap-A3
1155-1307 C1 DOMAIN
1308-1467 C2 DOMAIN

FIG. 4

10/33

HP44/OL NUCLEOTIDE SEQUENCE

1	ATGCAGCTAG	AGCTCTCCAC	CTGTGTCTTT	CTGTGTCTCT	TGCCACTCGG
	TACGTTCGATC	TCGAGAGGTG	GACACAGAAA	GACACAGAGA	ACGGTGAGCC
51	CTTTAGTGCC	ATCAGGAGAT	ACTACCTGGG	CGCAGTGGAA	CTGTCCTGGG
	GAAATCACGG	TAGTCCTCTA	TGATGGACCC	GCGTCACCTT	GACAGGACCC
101	ACTACCGGCA	AAGTGAACCTC	CTCCGTGAGC	TGCACGTGGA	CACCAGATTT
	TGATGGCCGT	TTCACCTTGAG	GAGGCACTCG	ACGTGCACCT	GTGGTCTAAA
151	CCTGCTACAG	CGCCAGGAGC	TCTTCCGTTG	GGCCCGTCAG	TCCTGTACAA
	GGACGATGTC	GCGGTCCTCG	AGAAGGCAAC	CCGGGCAGTC	AGGACATGTT
201	AAAGACTGTG	TTCGTAGAGT	TCACGGATCA	ACTTTTCAGC	GTTGCCAGGC
	TTTCTGACAC	AAGCATCTCA	AGTGCCTAGT	TGAAAAGTCG	CAACGGTCCG
251	CCAGGCCACC	ATGGATGGGT	CTGCTGGGTC	CTACCATCCA	GGCTGAGGTT
	GGTCCGGTGG	TACCTACCCA	GACGACCCAG	GATGGTAGGT	CCGACTCCAA
301	TACGACACGG	TGGTCGTTAC	CCTGAAGAAC	ATGGCTTCTC	ATCCCGTTAG
	ATGCTGTGCC	ACCAGCAATG	GGACTTCTTG	TACCGAAGAG	TAGGGCAATC
351	TCTTCACGCT	GTCGGCGTCT	CCTTCTGGAA	ATCTTCCGAA	GGCGCTGAAT
	AGAAGTGCGA	CAGCCGCAGA	GGAAGACCTT	TAGAAGGCTT	CCGCGACTTA
401	ATGAGGATCA	CACCAAGCCAA	AGGGAGAAGG	AAGACGATAA	AGTCCTTCCC
	TACTCCTAGT	GTGGTCGGTT	TCCCTCTTCC	TTCTGCTATT	TCAGGAAGGG
451	GGTAAAAGCC	AAACCTACGT	CTGGCAGGTC	CTGAAAAGAA	ATGGTCCAAC
	CCATTTTCGG	TTTGGATGCA	CACCGTCCAG	GACTTTCTTT	TACCAGGTTG
501	AGCCTCTGAC	CCACCATGTC	TTACCTACTC	ATACCTGTCT	CACGTGGACC
	TCGGAGACTG	GGTGGTACAG	AATGGATGAG	TATGGACAGA	GTGCACCTGG
551	TGGTGAAAGA	CCTGAATTCTG	GGCCTCATTG	GAGCCCTGCT	GGTTTGTAGA
	ACCACTTTCT	GGACTTAAGC	CCGGAGTAAC	CTCGGGACGA	CCAAACATCT
601	GAAGGGAGTC	TGACCAGAGA	AAGGACCCAG	AACCTGCACG	AATTTGTACT
	CTTCCCTCAG	ACTGGTCTCT	TTCCTGGGTC	TTGGACGTGC	TTAAACATGA
651	ACTTTTTTGCT	GTCTTTGATG	AAGGGAAAAG	TTGGCACTCA	GCAAGAAATG
	TGAAAAACGA	CAGAAACTAC	TTCCCTTTTC	AACCGTGAGT	CGTTCCTTAC
701	ACTCCTGGAC	ACGGGCCATG	GATCCCGCAC	CTGCCAGGGC	CCAGCCTGCA
	TGAGGACCTG	TGCCCCGTAC	CTAGGGCGTG	GACGGTCCCG	GGTCGGACGT
751	ATGCACACAG	TCAATGGCTA	TGTCAACAGG	TCTCTGCCAG	GTCTGATCGG
	TACGTGTGTC	AGTTACCGAT	ACAGTTGTCC	AGAGACGGTC	CAGACTAGCC
801	ATGTCATAAG	AAATCAGTCT	ACTGGCACGT	GATTGGAATG	GGCACCAGCC
	TACAGTATTC	TTTAGTCAGA	TGACCGTGCA	CTAACCTTAC	CCGTGGTCGG
851	CGGAAGTGCA	CTCCATTTTT	CTTGAAGGCC	ACACGTTTCT	CGTGAGGCAC
	GCCTTCACGT	GAGGTAAAAA	GAACCTCCGG	TGTGCAAAGA	GCACTCCGTG
901	CATCGCCAGG	CTTCCTTGGA	GATCTCGCCA	CTAACTTTCC	TCACTGCTCA
	GTAGCGGTCC	GAAGGAACCT	CTAGAGCGGT	GATTGAAAGG	AGTGACGAGT
951	GACATTCCTG	ATGGACCTTG	GCCAGTTCCT	ACTGTTTTGT	CATATCTCTT
	CTGTAAGGAC	TACCTGGAAC	CGGTCAAGGA	TGACAAAACA	GTATAGAGAA
1001	CCCACCACCA	TGGTGGCATG	GAGGCTCAGC	TCAGAGTAGA	AAGCTGCGCC
	GGGTGGTGGT	ACCACCGTAC	CTCCGAGTGC	AGTCTCATCT	TTTCGACGCG
1051	GAGGAGCCCC	AGCTGCGGAG	GAAAGCTGAT	GAAGAGGAAG	ATTATGATGA
	CTCCTCGGGG	TCGACGCCTC	CTTTCGACTA	CTTCTCCTTC	TAATACTACT
1101	CAATTTGTAC	GACTCGGACA	TGGACGTGGT	CCGGCTCGAT	GGTGACGACG
	GTTAAACATG	CTGAGCCTGT	ACCTGCACCA	GGCCGAGCTA	CCACTGCTGC
1151	TGTCTCCCTT	TATCCAAATC	CGCTCGGTTG	CCAAGAAGCA	TCCCAAACC
	ACAGAGGGAA	ATAGGTTTAG	GCGAGCCAAC	GGTTCTTCGT	AGGGTTTTGG
1201	TGGGTGCACT	ACATCTCTGC	AGAGGAGGAG	GACTGGGACT	ACGCCCCCGC
	ACCCACGTGA	TGTAGAGACG	TCTCCTCCTC	CTGACCCTGA	TGCGGGGGCG
1251	GGTCCCCAGC	CCCAGTGACA	GAAGTTATAA	AAGTCTCTAC	TTGAACAGTG
	CCAGGGGTCTG	GGGTCACTGT	CGGCAATATT	TTCAGAGATG	AACTTGTCAC
1301	GTCCTCAGCG	AATTGGTAGG	AAATACAAAA	AAGCTCGATT	CGTCGCTTAC
	CAGGAGTCGC	TTAACCATCC	TTTATGTTTT	TTCGAGCTAA	GCAGCGAATG

FIG. 5A

11/33

1351	ACGGATGTAA	CATTTAAGAC	TCGTAAAGCT	ATTCCGTATG	AATCAGGAAT
	TGCCTACATT	GTAAATTCTG	AGCATTTCTGA	TAAGGCATAC	TTAGTCCTTA
1401	CCTGGGACCT	TTACTTTTATG	GAGAAGTTGG	AGACACACTT	TTGATTATAT
	GGACCCTGGA	AATGAAATAC	CTCTTCAACC	TCTGTGTGAA	AACTAATATA
1451	TTAAGAATAA	AGCGAGCCGA	CCATATAACA	TCTACCCTCA	TGGAATCACT
	AATTCTTATT	TCGCTCGGCT	GGTATATTGT	AGATGGGAGT	ACCTTAGTGA
1501	GATGTCAGCG	CTTTGCACCC	AGGGAGACTT	CTAAAAGGTT	GGAAACATTT
	CTACAGTCGC	GAAACGTGGG	TCCCTCTGAA	GATTTTCCAA	CCTTTGTAAA
1551	GAAAGACATG	CCAATTCTGC	CAGGAGAGAC	TTTCAAGTAT	AAATGGACAG
	CTTTCTGTAC	GGTTAAGACG	GTCCTCTCTG	AAAGTTCATA	TTTACCTGTC
1601	TGACTGTGGA	AGATGGGCCA	ACCAAGTCCG	ATCCTCGGTG	CCTGACCCGC
	ACTGACACCT	TCTACCCGGT	TGGTTCAGGC	TAGGAGCCAC	GGACTGGGCG
1651	TACTACTCGA	GCTCCATTAA	TCTAGAGAAA	GATCTGGCTT	CGGGACTCAT
	ATGATGAGCT	CGAGGTAATT	AGATCTCTTT	CTAGACCGAA	GCCCTGAGTA
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAATC	TGTAGACCAA	AGAGGAAACC
	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTGGTT	TCTCCTTTGG
1751	AGATGATGTC	AGACAAGAGA	AACGTCATCC	TGTTTTCTGT	ATTTCGATGAG
	TCTACTACAG	TCTGTTCTCT	TTGCAGTAGG	ACAAAAGACA	TAAGCTACTC
1801	AATCAAAGCT	GGTACCTCGC	AGAGAATATT	CAGCGCTTCC	TCCCCAATCC
	TTAGTTTCGA	CCATGGAGCG	TCTCTTATAA	GTCGCGAAGG	AGGGGTTAGG
1851	GGATGGATTA	CAGCCCCAGG	ATCCAGAGTT	CCAAGCTTCT	AACATCATGC
	CCTACCTAAT	GTCGGGGTCC	TAGGTCTCAA	GGTTCGAAGA	TTGTAGTACG
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGCT	TGCAGCTGTC	GGTTTGTTTG
	TGTCGTAAGT	ACCGATACAA	AAACTATCGA	ACGTCGACAG	CCAAACAAAC
1951	CACGAGGTGG	CATACTGGTA	CATTCTAAGT	GTTGGAGCAC	AGACGGACTT
	GTGCTCCACC	GTATGACCAT	GTAAGATTCA	CAACCTCGTG	TCTGCCTGAA
2001	CCTCTCCGTC	TTCTTCTCTG	GCTACACCTT	CAAACACAAA	ATGGTCTATG
	GGAGAGGCAG	AAGAAGAGAC	CGATGTGGAA	GTTTGTGTTT	TACCAGATAC
2051	AAGACACACT	CACCCTGTTT	CCCTTCTCAG	GAGAAACGGT	CTTCATGTCA
	TTCTGTGTGA	GTGGGACAAG	GGGAAGAGTC	CTCTTTGCCA	GAAGTACAGT
2101	ATGGAAAACC	CAGGTCTCTG	GGTCCTTGGG	TGCCACAAC	CAGACTTGCG
	TACCTTTTGG	GTCCAGAGAC	CCAGGAACCC	ACGGTGTGTA	GTCTGAACGC
2151	GAACAGAGGG	ATGACAGCCT	TACTGAAGGT	GTATAGTTGT	GACAGGGACA
	CTTGTCTCCC	TACTGTCCGA	ATGACTTCCA	CATATCAACA	CTGTCCCTGT
2201	TTGGTGATTA	TTATGACAAC	ACTTATGAAG	ATATTCCAGG	CTTCTTGCTG
	AACCACTAAT	AATACTGTTG	TGAATACTTC	TATAAGGTCC	GAAGAACGAC
2251	AGTGGAAGA	ATGTCATTGA	ACCTAGGAGC	TTTGCCCAGA	ATTCAAGACC
	TCACCTTTCT	TACAGTAACT	TGGATCCTCG	AAACGGGTCT	TAAGTTCTGG
2301	CCCTAGTGCG	AGCGCTCCAA	AGCCTCCGGT	CCTGCGACGG	CATCAGAGGG
	GGGATCACGC	TCGCGAGGTT	TCGGAGGCCA	GGACGCTGCC	GTAGTCTCCC
2351	ACATAAGCCT	TCCTACTTTT	CAGCCGGAGG	AAGACAAAAT	GGACTATGAT
	TGTATTTCGA	AGGATGAAAA	GTCGGCCTCC	TTCTGTTTTA	CCTGATACTA
2401	GATATCTTCT	CAACTGAAAC	GAAGGGAGAA	GATTTTGACA	TTTACGGTGA
	CTATAGAAGA	GTTGACTTTG	CTTCCCTCTT	CTAAAACGTG	AAATGCCACT
2451	GGATGAAAAT	CAGGACCCTC	GCAGCTTTCA	GAAGAGAACC	CGACACTATT
	CCTACTTTTA	GTCCTGGGAG	CGTCGAAAGT	CTTCTCTTGG	GCTGTGATAA
2501	TCATTGCTGC	GGTGGAGCAG	CTCTGGGATT	ACGGGATGAG	CGAATCCCCC
	AGTAACGACG	CCACCTCGTC	GAGACCCTAA	TGCCCTACTC	GCTTAGGGGG
2551	CGGGCGCTAA	GAAACAGGGC	TCAGAACGGA	GAGGTGCCTC	GGTTCAAGAA
	GCCCGCGATT	CTTTGTCCCC	AGTCTTGCCCT	CTCCACGGAG	CCAAGTTCTT
2601	GGTGGTCTTC	CGGGAATTTG	CTGACGGCTC	CTTCACGCAG	CCGTCGTACC
	CCACCAGAAG	GCCCTTAAAC	GAAGTGCCTC	GAAGTGCCTC	GGCAGCATGG
2651	GCGGGGAACT	CAACAAACAC	TTGGGGCTCT	TGGGACCCTA	CATCAGAGCG
	CGCCCTTGA	GTTGTTTGTG	AACCCCGAGA	ACCCTGGGAT	GTAGTCTCGC
2701	GAAGTTGAAG	ACAACATCAT	GGTAACTTTC	AAAAACCAGG	CGTCTCGTCC
	CTTCAACTTC	TGTTGTAGTA	CCATTGAAAG	TTTTTGGTCC	GCAGAGCAGG
2751	CTATTCCTTC	TACTCGAGCC	TTATTTCTTA	TCCGGATGAT	CAGGAGCAAG

FIG. 5B

12/33

	GATAAGGAAG	ATGAGCTCGG	AATAAAGAAT	AGGCCTACTA	GTCCTCGTTC
2801	GGGCAGAACC	TCGACACAAC	TTCGTCCAGC	CAAATGAAAC	CAGAACTTAC
	CCCGTCTTGG	AGCTGTGTTG	AAGCAGGTCG	GTTTACTTTG	GTCTTGAATG
2851	TTTTGGAAAG	TGCAGCATCA	CATGGCACCC	ACAGAAGACG	AGTTTGA CTG
	AAAACCTTTC	ACGTCGTA GT	GTACCGTGGG	TGTCTTCTGC	TCAAAC TGAC
2901	CAAAGCCTGG	GCCTACTTTT	CTGATGTTGA	CCTGGAAAAA	GATGTGCACT
	GTTTCGGACC	CGGATGAAAA	GACTACA ACT	GGACCTTTTT	CTACACGTGA
2951	CAGGCTTGAT	CGGCCCCCTT	CTGATCTGCC	GCGCCAACAC	CCTGAACGCT
	GTCCGA ACTA	GCCGGGGGAA	GACTAGACGG	CGCGGTGTG	GGACTTGCGA
3001	GCTCACGGTA	GACAAGTGAC	CGTGCAAGAA	TTTGCTCTGT	TTTTCACTAT
	CGAGTGCCAT	CTGTTCACTG	GACCGTTCTT	AAACGAGACA	AAAAGTGATA
3051	TTTTGATGAG	ACAAAGAGCT	GGTACTTCAC	TGAAAATGTG	GAAAGGAACT
	AAAAC TACTC	TGTTTCTCGA	CCATGAAGTG	ACTTTTACAC	CTTTCCTTGA
3101	GCCGGGCCCC	CTGCCATCTG	CAGATGGAGG	ACCCCACTCT	GAAAGAAAAC
	CGGCCCCGGG	GACGGTAGAC	GTCTACCTCC	TGGGGTGAGA	CTTTCCTTTG
3151	TATCGCTTCC	ATGCAATCAA	TGGCTATGTG	ATGGATACAC	TCCCTGGCTT
	ATAGCGAAGG	TACGTTAGTT	ACCGATACAC	TACCTATGTG	AGGGACCGAA
3201	AGTAATGGCT	CAGAATCAAA	GGATCCGATG	GTATCTGCTC	AGCATGGGCA
	TCATTACCGA	GTCTTAGTTT	CCTAGGCTAC	CATAGACGAG	TCGTACCCGT
3251	GCAATGAAAA	TATCCATTTCG	ATTCA TTTTA	GCGGACACGT	GTTCAGTGTA
	CGTTACTTTT	ATAGGTAAGC	TAAGTAAAAT	CGCCTGTGCA	CAAGTCACAT
3301	CGGAAAAAGG	AGGAGTATAA	AATGGCCGTG	TACAATCTCT	ATCCGGGTGT
	GCCTTTT TCC	TCCTCATATT	TTACCGGCAC	ATGTTAGAGA	TAGGCCCACA
3351	CTTTGAGACA	GTGGAAATGC	TACCGTCCAA	AGTTGGAATT	TGGCGAATAG
	GAAACTCTGT	CACCTTTACG	ATGGCAGGTT	TCAACCTTAA	ACCGCTTATC
3401	AATGCCTGAT	TGGCGAGCAC	CTGCAAGCTG	GGATGAGCAC	GACTTTCCTG
	TTACGGACTA	ACCGCTCGTG	GACGTT CGAC	CCTACTCGTG	CTGAAAGGAC
3451	GTGTACAGCA	AGAAGTGTCA	GACTCCCCTG	GGAATGGCTT	CTGGACACAT
	CACATGTCTG	TCTTCACAGT	CTGAGGGGAC	CCTTACCGAA	GACCTGTGTA
3501	TAGAGATTTT	CAGATTACAG	CTTCAGGACA	ATATGGACAG	TGGGCCCCAA
	ATCTCTAAAA	GTCTAATGTC	GAAGTCCTGT	TATACCTGTC	ACCCGGGGTT
3551	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA	TCAATGCCTG	GAGCACCAAG
	TCGACCGGTC	TGAAGTAATA	AGGCCTAGTT	AGTTACGGAC	CTCGTGGTTC
3601	GAGCCCTTTT	CTTGGA TCAA	GGTGGA TCTG	TTGGCACC AA	TGATTATTCA
	CTCGGGAAAA	GAACCTAGTT	CCACCTAGAC	AACCGTGGTT	ACTAATAAGT
3651	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA	GTTCTCCAGC	CTCTACATCT
	GCCGTAGTTC	TGGGTCCCAC	GGGCAGTCTT	CAAGAGGTCC	GAGATGTAGA
3701	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG	GGAAGAAGTG	GCAGACTTAT
	GAGTCAAATA	GTAGTACATA	TCAGAACTAC	CCTTCTTCAC	CGTCTGAATA
3751	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC	TTCTTTGGCA	ATGTGGATT C
	GCTCCTTTAA	GGTGACCTTG	GAATTACCAG	AAGAAACCGT	TACACCTAAG
3801	ATCTGGGATA	AAACACAATA	TTTTTAAACC	TCCAATTATT	GCTCGATACA
	TAGACCCTAT	TTTGTGTTAT	AAAAATTGGG	AGGTTAATAA	CGAGCTATGT
3851	TCCGTTTGCA	CCCAACTCAT	TATAGCATTC	GCAGCACTCT	TCGCATGGAG
	AGGCAAACGT	GGGTTGAGTA	ATATCGTAAG	CGTCGTCAGA	AGCGTACCTC
3901	TTGATGGGCT	GTGATTTAAA	TAGTTGCAGC	ATGCCATTGG	GAATGGAGAG
	AACTACCCGA	CACTAAATTT	ATCAACGTCG	TACGGTAACC	CTTACCTCTC
3951	TAAAGCAATA	TCAGATGCAC	AGATTACTGC	TTCATCCTAC	TTTACCAATA
	ATTTTCGTTAT	AGTCTACGTG	TCTAATGACG	AAGTAGGATG	AAATGGTTAT
4001	TGTTTGCCAC	CTGGTCTCCT	TCAAAAGCTC	GACTTCACCT	CCAAGGGAGG
	ACAAACGGTG	GACCAGAGGA	AGTTTTTCGAG	CTGAAGTGGA	GGTTCCCTCC
4051	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT	CCAAAAGAGT	GGCTGCAAGT
	TCATTACGGA	CCTCTGGAGT	CCACTTATTA	GGTTTTCTCA	CCGACGTTCA
4101	GGACTTCCAG	AAGACAATGA	AAGTCACAGG	AGTAACTACT	CAGGGAGTAA
	CCTGAAGGTC	TTCTGTTACT	TTCAGTGTCC	TCATTGATGA	GTCCCTCATT
4151	AATCTCTGCT	TACCAGCATG	TATGTGAAGG	AGTTCCTCAT	CTCCAGCAGT
	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC	TCAAGGAGTA	GAGGTCGTCA

FIG. 5C

13/33

```
4201 CAAGATGGCC ATCAGTGGAC TCTCTTTTTT CAGAATGGCA AAGTAAAGGT
      GTTCTACCGG TAGTCACCTG AGAGAAAAAA GTCTTACCGT TTCATTTCCA
4251 TTTTCAGGGA AATCAAGACT CCTTCACACC TGTGGTGAAC TCTCTAGACC
      AAAAGTCCCT TTAGTTCTGA GGAAGTGTGG ACACCACTTG AGAGATCTGG
4301 CACCGTTACT GACTCGCTAC CTTCGAATTC ACCCCCAGAG TTGGGTGCAC
      GTGGCAATGA CTGAGCGATG GAAGCTTAAG TGGGGGTCTC AACCACGTG
4351 CAGATTGCCC TGAGGATGGA GGTTCTGGGC TGCAGGCAC AGGACCTCTA
      GTCTAACGGG ACTCCTACCT CCAAGACCCG ACGCTCCGTG TCCTGGAGAT
4401 C
      G
```

1-57 SIGNAL PEPTIDE
58-1173 A1 DOMAIN
1174-2277 A2 DOMAIN
2278-2349 OL LINKER
2350-3462 ap-A3 DOMAINS
3463-3921 C1 DOMAIN
3922-4401 C2 DOMAIN

FIG. 5D

14/33

AMINO ACID SEQUENCE OF HP46/SQ

```
1  MOLELSTCVF LCLLPLGFSA IRRYYLGAVE LSWDYRQSEL LRELHVDTRF
51  PATAPGALPL GPSVLYKKTV FVEFTDQLFS VARPRPPWMG LLGPTIQAEV
101 YDTVVVTLKN MASHPVSLHA VGVSFWKSSS GAIEDHTSQ REKEDDKVLP
151 GKSQTYVWQV LKENGPTASD PPCLTYSYLS HVDLVKDLNS GLIGALLVCR
201 EGSLTRERTQ NLHEFVLLFA VFDEGKSWHS ARNDSWTRAM DPAPARAQPA
251 MHTVNGYVNR SLPGLIGCHK KSVYWHVIGM GTSPEVHSIF LEGHTFLVRH
301 HRQASLEISP LTFLTAQTFL MDLGQFLLFC HISSHHHGGM EAHVRVESCA
351 EEPQLRRKAD EEEDYDDNLY DSDMDVVRDL GDDVSPFIQI RSVAKKHPKT
401 WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQIRGR KYKKVRFMAY
451 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR PYNIYPHGIT
501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP TKSDPRCLTR
551 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR NVILFSVFDE
601 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV FDSLQLSVCL
651 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEDTLTLF PFSGETVFMS
701 MENPGLWILG CHNSFLNRG MTALLKVSSC DKNTGDYED SYEDISAYLL
751 SKNNAIEPRS FSQNPPVLKR HQREITRRTL QSDQEEIDYD DTISVEMKKE
801 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP HVLNRNAQSG
851 SVPQFKKVVV QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA EVEDNIMVTF
901 RNQASRPYSF YSSLISYED QROGAEPKRN FVKPNETKTY FWKVQHMAP
951 TKDEFDCAW AYFSDVDLEK DVHSGLIGPL LVCHTNLTNP AHGRQVTQVE
1001 FALFFTIFDE TKSIFYTENM ERNCRAPCNI QMEDPTFKEN YRFHAINGYI
1051 MDTLPGLVMA QDQIRWYLL SMGSNENIHS IHFSGHVFTV RKKEEYKMAI
1101 YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL VYSNKCQTPL
1151 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK EPFSWIKVDL
1201 LAPMIIHGK TQGARQKFSS LYISQFIIMY SLDGKKWQTY RGNSTGTLMV
1251 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME LMGCDLNSCS
1301 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR SNAWRPQVNN
1351 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISS QDGHQWTLFF
1401 QNGKVVFQNG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH QIALRMEVLG
1451 CEAQDLY*
```

1-19 SIGNAL PEPTIDE
20-391 A1 DOMAIN
392-759 A2 DOMAIN
760-773 SQ LINKER
774-1144 ap-A3
1145-1297 C1 DOMAIN
1298-1457 C2 DOMAIN

FIG. 6

15/33

HP46/SQ NUCLEOTIDE SEQUENCE

1	ATGCAGCTAG	AGCTCTCCAC	CTGTGTCTTT	CTGTGTCTCT	TGCCACTCGG
	TACGTCGATC	TCGAGAGGTG	GACACAGAAA	GACACAGAGA	ACGGTGAGCC
51	CTTTAGTGCC	ATCAGGAGAT	ACTACCTGGG	CGCAGTGGA	CTGTCCTGGG
	GAAATCACGG	TAGTCCTCTA	TGATGGACCC	GCGTCACCTT	GACAGGACCC
101	ACTACCGGCA	AAGTGAACCTC	CTCCGTGAGC	TGCACGTGGA	CACCAGATTT
	TGATGGCCGT	TTCACCTTGAG	GAGGCACTCG	ACGTGCACCT	GTGGTCTAAA
151	CCTGCTACAG	CGCCAGGAGC	TCTTCCGTTG	GGCCCGTCAG	TCCTGTACAA
	GGACGATGTC	GCGGTCCTCG	AGAAGGCAAC	CCGGGCAGTC	AGGACATGTT
201	AAAGACTGTG	TTCGTAGAGT	TCACGGATCA	ACTTTTTCAGC	GTTGCCAGGC
	TTTCTGACAC	AAGCATCTCA	AGTGCCTAGT	TGAAAAGTCG	CAACGGTCCG
251	CCAGGCCACC	ATGGATGGGT	CTGCTGGGTC	CTACCATCCA	GGCTGAGGTT
	GGTCCGGTGG	TACCTACCCA	GACGACCCAG	GATGGTAGGT	CCGACTCCAA
301	TACGACACGG	TGGTCGTTAC	CCTGAAGAAC	ATGGCTTCTC	ATCCCGTTAG
	ATGCTGTGCC	ACCAGCAATG	GGACTTCTTG	TACCGAAGAG	TAGGGCAATC
351	TCTTCACGCT	GTCGGCGTCT	CCTTCTGGAA	ATCTTCCGAA	GGCGCTGAAT
	AGAAGTGCGA	CAGCCGCAGA	GGAAGACCTT	TAGAAGGCTT	CCGCGACTTA
401	ATGAGGATCA	CACCAGCCAA	AGGGAGAAGG	AAGACGATAA	AGTCCTTCCC
	TACTCCTAGT	GTGGTCGGTT	TCCCTCTTCC	TTCTGCTATT	TCAGGAAGGG
451	GGTAAAAGCC	AAACCTACGT	CTGGCAGGTC	CTGAAAGAAA	ATGGTCCAAC
	CCATTTTCGG	TTTGGATGCA	GACCGTCCAG	GACTTTTCTTT	TACCAGGTTG
501	AGCCTCTGAC	CCACCATGTC	TTACCTACTC	ATACCTGTCT	CACGTGGACC
	TCGGAGACTG	GGTGGTACAG	AATGGATGAG	TATGGACAGA	GTGCACCTGG
551	TGGTGAAAGA	CCTGAATTCTG	GGCCTCATTG	GAGCCCTGCT	GGTTTGTAGA
	ACCACTTTCT	GGACTTAAGC	CCGGAGTAAC	CTCGGGACGA	CCAAACATCT
601	GAAGGGAGTC	TGACCAGAGA	AAGGACCCAG	AACCTGCACG	AATTTGTACT
	CTTCCCTCAG	ACTGGTCTCT	TTCTTGGGTC	TTGGACGTGC	TTAAACATGA
651	ACTTTTGTCT	GTCTTTGATG	AAGGGAAAAG	TTGGCACTCA	GCAAGAAATG
	TGAAAAACGA	CAGAAACTAC	TTCCCTTTTC	AACCGTGAGT	CGTCTTTTAC
701	ACTCCTGGAC	ACGGGCCATG	GATCCCGCAC	CTGCCAGGGC	CCAGCCTGCA
	TGAGGACCTG	TGCCCCGTAC	CTAGGGCGTG	GACGGTCCCCG	GGTCGGACGT
751	ATGCACACAG	TCAATGGCTA	TGTCAACAGG	TCTCTGCCAG	GTCTGATCGG
	TACGTGTGTC	AGTTACCGAT	ACAGTTGTCC	AGAGACGGTC	CAGACTAGCC
801	ATGTCATAAG	AAATCAGTCT	ACTGGCACGT	GATTGGAATG	GGCACCAGCC
	TACAGTATTC	TTTAGTCAGA	TGACCGTGCA	CTAACCTTAC	CCGTGGTCGG
851	CGGAAGTGCA	CTCCATTTTTT	CTTGAAGGCC	ACACGTTTCT	CGTGAGGCAC
	GCCTTCACGT	GAGGTAAAAA	GAAGTTCCGG	TGTGCAAAGA	GCACTCCGTG
901	CATCGCCAGG	CTTCCTTGGA	GATCTCGCCA	CTAACTTTCC	TCACTGCTCA
	GTAGCGGTCC	GAAGGAACCT	CTAGAGCGGT	GATTGAAAGG	AGTGACGAGT
951	GACATTCCTG	ATGGACCTTG	GCCAGTTCCT	ACTGTTTTGT	CATATCTCTT
	CTGTAAGGAC	TACCTGGAAC	CGGTCAAGGA	TGACAAAACA	GTATAGAGAA
1001	CCCACCACCA	TGGTGGCATG	GAGGCTCACG	TCAGAGTAGA	AAGCTGCGCC
	GGGTGGTGGT	ACCACCGTAC	CTCCGAGTGC	AGTCTCATCT	TTCGACGCGG
1051	GAGGAGCCCC	AGCTGCGGAG	GAAAGCTGAT	GAAGAGGAAG	ATTATGATGA
	CTCCTCGGGG	TCGACGCCTC	CTTTCGACTA	CTTCTCCTTC	TAATACTACT
1101	CAATTTGTAC	GACTCGGACA	TGGACGTGGT	CCGGCTCGAT	GGTGACGACG
	GTAAACATG	CTGAGCCTGT	ACCTGCACCA	GGCCGAGCTA	CCACTGCTGC
1151	TGTCTCCCTT	TATCCAAATC	CGCTCAGTTG	CCAAGAAGCA	TCCTAAAACT
	ACAGAGGGAA	ATAGGTTTAT	GCGAGTCAAC	GGTTCTTCGT	AGGATTTTGA
1201	TGGGTACATT	ACATTGCTGC	TGAAGAGGAG	GACTGGGACT	ATGCTCCCTT

FIG. 7A

16/33

1251	ACCCATGTAA	TGTAACGACG	ACTTCTCCTC	CTGACCCTGA	TACGAGGGAA
	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAATG
	TCAGGAGCGG	GGGCTACTGT	CTTCAATATT	TTCAGTTATA	AACTTGTTAC
1301	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAA	AAGTCCGATT	TATGGCATACT
	CGGGAGTCGC	CTAACCATCC	TTCATGTTTT	TTCAGGCTAA	ATACCGTATG
1351	ACAGATGAAA	CCTTTAAGAC	GCGTGAAGCT	ATTCAGCATG	AATCAGGAAT
	TGTCTACTTT	GGAAATTCTG	CGCACTTCGA	TAAGTCGTAC	TTAGTCCTTA
1401	CTTGGGACCT	TTACTTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
	GAACCCTGGA	AATGAAATAC	CCCTTCAACC	TCTGTGTGAC	AACTAATATA
1451	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	TCTACCCTCA	CGGAATCACT
	AATTCTTAGT	TCGTTCGTCT	GGTATATTGT	AGATGGGAGT	GCCTTAGTGA
1501	GATGTCCGTC	CTTTGTATTC	AAGGAGATTA	CCAAAAGGTG	TAAAACATTT
	CTACAGGCAG	GAAACATAAG	TTCTCTAAT	GGTTTTCCAC	ATTTTGTAAA
1551	GAAGGATTTT	CCAATTCTGC	CAGGAGAAAT	ATTCAAATAT	AAATGGACAG
	CTTCCTAAAA	GGTTAAGACG	GTCCTCTTTA	TAAGTTTATA	TTTACCTGTC
1601	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCGCGGGT	CCTGACCCGC
	ACTGACATCT	TCTACCCGGT	TGATTTAGTC	TAGGCGCCAC	GGACTGGGCG
1651	TATTACTCTA	GTTTCGTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
	ATAATGAGAT	CAAAGCAATT	ATACCTCTCT	CTAGATCGAA	GTCCTGAGTA
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAATC	TGTAGTCCAA	AGAGGAAACC
	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTAGTT	TCTCCTTTGG
1751	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTTCTGT	ATTTGATGAG
	TCTATTACAG	TCTGTTCTCC	TTACAGTAGG	ACAAAAGACA	TAAACTACTC
1801	AACCGAAGCT	GGTACCTCAC	AGAGAATATA	CAACGCTTTC	TCCCCAATCC
	TTGGCTTCGA	CCATGGAGTG	TCTCTTATAT	GTTGCGAAAG	AGGGGTTAGG
1851	AGCTGGAGTA	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
	TCGACCTCAC	GTCGAACTCC	TAGGTCTCAA	GGTTCGAAGG	TTGTAGTACG
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTTGTTTG
	TGTCGTAGTT	ACCGATACAA	AAACTATCAA	ACGTCAACAG	TCAAACAAAC
1951	CATGAGGTGG	CATACTGGTA	CATTCTAAGC	ATTGGAGCAC	AGACTGACTT
	GTATGCCACC	GTATGACCAT	GTAAGATTCCG	TAACCTCGTG	TCTGACTGAA
2001	CCTTTCTGTC	TTCTTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
	GGAAAGACAG	AAGAAGAGAC	CTATATGGAA	GTTTGTGTTT	TACCAGATAC
2051	AAGACACACT	CACCCTATTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTCTG
	TTCTGTGTGA	GTGGGATAAG	GGTAAGAGTC	CTCTTTGACA	GAAGTACAGC
2101	ATGGAAAACC	CAGGTCTATG	GATTCTGGGG	TGCCACAACCT	CAGACTTTTCG
	TACCTTTTGG	GTCCAGATAC	CTAAGACCCC	ACGGTGTGTA	GTCTGAAAGC
2151	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAACA
	CTTGTCTCCG	TACTGGCGGA	ATGACTTCCA	AAGATCAACA	CTGTTCTTGT
2201	CTGGTGATTA	TTACGAGGAC	AGTTATGAAG	ATATTTTCAGC	ATACTTGCTG
	GACCACTAAT	AATGCTCCTG	TCAATACTTC	TATAAAGTCG	TATGAACGAC
2251	AGTAAAAACA	ATGCCATTGA	ACCTAGGAGC	TTCTCTCAGA	ATCCACCAGT
	TCATTTTTGT	TACGGTAACT	TGGATCCTCG	AAGAGAGTCT	TAGGTGGTCA
2301	CTTGAAACGC	CATCAACGGG	AAATAACTCG	TACTACTCTT	CAGTCAGATC
	GAACCTTTCG	GTAAGTTGCC	TTTATTGAGC	ATGATGAGAA	GTCAGTCTAG
2351	AAGAGGAAAT	TGACTATGAT	GATACCATAT	CAGTTGAAAT	GAAGAAGGAA
	TTCTCCTTTA	ACTGATACTA	CTATGGTATA	GTCAACTTTA	CTTCTTCCTT
2401	GATTTTGTGA	TTTATGATGA	GGATGAAAAT	CAGAGCCCCC	GCAGCTTTCA
	CTAAACTGT	AAATACTACT	CCTACTTTTA	GTCTCGGGGG	CGTCGAAAGT
2451	AAAGAAAACA	CGACACTATT	TTATTGCTGC	AGTGGAGAGG	CTCTGGGATT
	TTTCTTTTGT	GCTGTGATAA	AATAACGACG	TCACCTCTCC	GAGACCCTAA
2501	ATGGGATGAG	TAGCTCCCCA	CATGTTCTAA	GAAACAGGGC	TCAGAGTGGC

FIG. 7B

17/33

2551	TACCCTACTC	ATCGAGGGGT	GTACAAGATT	CTTTGTCCCC	AGTCTCAGCG
	AGTGTCCCTC	AGTTCAAGAA	AGTTGTTTTT	CAGGAATTTA	CTGATGGCTC
	TCACAGGGAG	TCAAGTTCTT	TCAACAAAAG	GTCCTTAAAT	GACTACCGAG
2601	CTTTACTCAG	CCCTTATACC	GTGGAGAACT	AAATGAACAT	TTGGGACTCC
	GAAATGAGTC	GGGAATATGG	CACCTCTTGA	TTTACTTGTA	AACCCCTGAGG
2651	TGGGGCCATA	TATAAGAGCA	GAAGTTGAAG	ATAATATCAT	GGTAACTTTC
	ACCCCGGTAT	ATATTCTCGT	CTTCAACTTC	TATTATAGTA	CCATTGAAAG
2701	AGAAATCAGG	CCTCTCGTCC	CTATTCCTTC	TATTCTAGCC	TTATTTCTTA
	TCTTTAGTCC	GGAGAGCAGG	GATAAGGAAG	ATAAGATCGG	AATAAAGAAT
2751	TGAGGAAGAT	CAGAGGCAAG	GAGCAGAACC	TAGAAAAAAC	TTTGTCAAGC
	ACTCCTTCTA	GTCTCCGTTC	CTCGTCTTGG	ATCTTTTTTG	AAACAGTTCC
2801	CTAATGAAAC	CAAACTTAC	TTTTGGAAAG	TGCAACATCA	TATGGCACCC
	GATTACTTTG	GTTTTGAATG	AAAACCTTTC	ACGTTGTAGT	ATACCGTGGG
2851	ACTAAAGATG	AGTTTGACTG	CAAAGCCTGG	GCTTATTTCT	CTGATGTTGA
	TGATTTCTAC	TCAAACGAC	GTTTCGGACC	CGAATAAAGA	GACTACAAC
2901	CCTGGAAAAA	GATGTGCACT	CAGGCCTGAT	TGGACCCCTT	CTGGTCTGCC
	GGACCTTTTT	CTACACGTGA	GTCCGGACTA	ACCTGGGGAA	GACCAGACGG
2951	ACACTAACAC	ACTGAACCCT	GCTCATGGGA	GACAAGTGAC	AGTACAGGAA
	TGTGATTGTG	TGACTTGGGA	CGAGTACCCT	CTGTTCAC	TCATGTCATT
3001	TTTGCTCTGT	TTTTCACCAT	CTTTGATGAG	ACCAAAAGCT	GGTACTTCAC
	AAACGAGACA	AAAAGTGGA	GAAACTACTC	TGGTTTTCGA	CCATGAAGTG
3051	TGAAAATATG	GAAAGAACT	GCAGGGCTCC	CTGCAATATC	CAGATGGAAG
	ACTTTTATAC	CTTTCTTTGA	CGTCCCGAGG	GACGTTATAG	GTCTACCTTC
3101	ATCCCACTTT	TAAAGAGAAT	TATCGCTTCC	ATGCAATCAA	TGGCTACATA
	TAGGGTGAAA	ATTTCTCTTA	ATAGCGAAGG	TACGTTAGTT	ACCGATGTAT
3151	ATGGATACAC	TACCTGGCTT	AGTAATGGCT	CAGGATCAAA	GGATTCCGATG
	TACCTATGTG	ATGGACCGAA	TCATTACCGA	GTCCTAGTTT	CCTAAGCTAC
3201	GTATCTGCTC	AGCATGGGCA	GCAATGAAAA	CATCCATTCT	ATTCAATTTCA
	CATAGACGAG	TCGTACCCGT	CGTTACTTTT	GTAGGTAAAG	TAAAGTAAAGT
3251	GTGGACATGT	GTTCACTGTA	GAAAAAAAG	AGGAGTATAA	AATGGCACTG
	CACCTGTACA	CAAGTGACAT	GCTTTTTTTC	TCCTCATATT	TTACCGTGAC
3301	TACAATCTCT	ATCCAGGTGT	TTTTGAGACA	GTGGAAATGT	TACCATCCAA
	ATGTTAGAGA	TAGGTCCACA	AAAACCTCTG	CACCTTTACA	ATGGTAGGTT
3351	AGCTGGAATT	TGGCGGGTGG	AATGCCTTAT	TGGCGAGCAT	CTACATGCTG
	TCGACCTTAA	ACCGCCCACC	TTACGGAATA	ACCGCTCGTA	GATGTACGAC
3401	GGATGAGCAC	ACTTTTTCTG	GTGTACAGCA	ATAAGTGTCA	GACTCCCCTG
	CCTACTCGTG	TGAAAAAGAC	CACATGTCGT	TATTCACAGT	CTGAGGGGAC
3451	GGAATGGCTT	CTGGACACAT	TAGAGATTTT	CAGATTACAG	CTTCAGGACA
	CCTTACCGAA	GACCTGTGTA	ATCTCTAAAA	GTCTAATGTC	GAAGTCCTGT
3501	ATATGGACAG	TGGGCCCCAA	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA
	TATACCTGTC	ACCCGGGGTT	TCGACCGGTC	TGAAGTAATA	AGGCCTAGTT
3551	TCAATGCCTG	GAGCACCAAG	GAGCCCTTTT	CTTGGATCAA	GGTGGATCTG
	AGTTACGGAC	CTCGTGGTTC	CTCGGGAAAA	GAACCTAGTT	CCACCTAGAC
3601	TTGGCACCAA	TGATTATTCA	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA
	AACCGTGGTT	GCCGTAGTTC	GCCGTAGTTC	TGGGTCCAC	GGGCAGTCTT
3651	GTTCTCCAGC	CTCTACATCT	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG
	CAAGAGGTCG	GAGATGTAGA	GAGTCAAATA	GTAGTACATA	TCAGAACTAC
3701	GGAAGAAGTG	GCAGACTTAT	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC
	CCTTCTTCAC	CGTCTGAATA	GCTCCTTTAA	GGTGACCTTG	GAATTACCAG
3751	TTCTTTGGCA	ATGTGGATTTC	ATCTGGGATA	AAACACAATA	TTTTTAACCC
	AAGAAACCGT	TACACCTAAG	TAGACCCTAT	TTTGTGTTAT	AAAAATTGGG
3801	TCCAATTATT	GCTCGATACA	TCCGTTTGCA	CCCAACTCAT	TATAGCATTC

FIG. 7C

18/33

	AGGTTAATAA	CGAGCTATGT	AGGCAAAGCT	GGGTTGAGTA	ATATCGTAAG
3851	GCAGCACTCT	TCGCATGGAG	TTGATGGGCT	GTGATTTAAA	TAGTTGCAGC
	CGTCGTGAGA	AGCGTACCTC	AACTACCCGA	CACTAAATTT	ATCAACGTCC
3901	ATGCCATTGG	GAATGGAGAG	TAAAGCAATA	TCAGATGCAC	AGATTACTGC
	TACGGTAACC	CTTACCTCTC	ATTTTCGTTAT	AGTCTACGTG	TCTAATGACG
3951	TTCATCCTAC	TTTACCAATA	TGTTTGCCAC	CTGGTCTCCT	TCAAAAGCTC
	AAGTAGGATG	AAATGGTTAT	ACAAACGGTG	GACCAGAGGA	AGTTTTTCGAG
4001	GACTTCACCT	CCAAGGGAGG	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT
	CTGAAGTGGA	GGTTCCCTCC	TCATTACGGA	CCTCTGGAGT	CCACTTATTA
4051	CCAAAAGAGT	GGCTGCAAGT	GGACTTCCAG	AAGACAATGA	AAGTCACAGG
	GGTTTTCTCA	CCGACGTTCA	CCTGAAGGTC	TTCTGTTACT	TTCAGTGTCC
4101	AGTAACTACT	CAGGGAGTAA	AATCTCTGCT	TACCAGCATG	TATGTGAAGG
	TCATTGATGA	GTCCCTCAT	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC
4151	AGTTCCTCAT	CTCCAGCAGT	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTTT
	TCAAGGAGTA	GAGGTCGTCA	GTTCTACCGG	TAGTCACCTG	AGAGAAAAAA
4201	CAGAATGGCA	AAGTAAAGGT	TTTTTCAGGA	AATCAAGACT	CCTTCACACC
	GTCTTACCGT	TTCATTTCCA	AAAACCTCCCT	TTAGTTCTGA	GGAAGTGTGG
4251	TGTGGTGAAC	TCTCTAGACC	CACCGTTACT	GACTCGCTAC	CTTCGAATTC
	ACACCACTTG	AGAGATCTGG	GTGGCAATGA	CTGAGCGATG	GAAGCTTAAG
4301	ACCCCCAGAG	TTGGGTGCAC	CAGATTGCCC	TGAGGATGGA	GGTTCTGGGC
	TGGGGGTCTC	AACCCACGTG	GTCTAACGGG	ACTCCTACCT	CCAAGACCCG
4351	TGCGAGGCAC	AGGACCTCTA	C		
	ACGCTCCGTG	TCCTGGAGAT	G		

1-57 SIGNAL PEPTIDE
58-1173 A1 DOMAIN
1174-2277 A2 DOMAIN
2278-2319 SQ LINKER
2320-3432 ap-A3 DOMAINS
3433-3891 C1 DOMAIN
3892-4371 C2 DOMAIN

FIG. 7D

19/33

AMINO ACID SEQUENCE OF HP47/OL

1	MQLELSTCVF	LCLLPLGFSA	IRRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYVWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLFC	HISHHHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRD	GDDVSPFIQI	RSVAKKHPKT
401	WVHYIAAEEE	DWDYAPLVLA	PDDRSYKSQY	LNNGPQRIGR	KYKKVRFMAY
451	TDETFTKTREA	IQHESGILGP	LLYGEVGDTL	LIIFKNQASR	PYNIYPHGIT
501	DVRPLYSRRL	PKGVKHLKDF	PILPGEIFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSFVNMER	DLASGLIGPL	LICYKESVDQ	RGNQIMSDKR	NVILFSVFDE
601	NRSWYL TENI	QRFLPNPAGV	QLEDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYILS	IGAQTDFLSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWILG	CHNSDFRNRG	MTALLKVSSC	DKNTGDYED	SYEDISAYLL
751	SKNNAIEPRS	FAQNSRPPSA	SAPKPPVLR	HORDISLPTF	QPEEDKMDYD
801	DIFSTETKGE	DFDIYGEDEN	QDPRSFOKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALNR AQNG	EVPRFKKVVF	REFADGSFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVTF	KNQASRPYSF	YSSLISYPDD	QEOGAEPHRN	FVQPNETRTY
951	FWKVQHMAP	TEDEFDCKAW	AYFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLKEN
1051	YRFHAINGYV	MDTLPGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFSV
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVGI	WRIECLIGEH	LQAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASGOYGO	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKVDL	LAPMIIHGK	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNIFNPPII	ARYIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLQGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISSS
1401	QDGHQWTLFF	QNGKVKVFOG	NQDSFTPVVN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY*			

1-19 SIGNAL PEPTIDE
20-391 A1 DOMAIN
392-759 A2 DOMAIN
760-783 OL LINKER
784-1154 ap-A3
1155-1307 C1 DOMAIN
1308-1467 C2 DOMAIN

FIG. 8

20/33

HP47/OL NUCLEOTIDE SEQUENCE

1	ATGCAGCTAG	AGCTCTCCAC	CTGTGTCTTT	CTGTGTCTCT	TGCCACTCGG
	TACGTCGATC	TCGAGAGGTG	GACACAGAAA	GACACAGAGA	ACGGTGAGCC
51	CTTTAGTGCC	ATCAGGAGAT	ACTACCTGGG	CGCAGTGGA	CTGTCCTGGG
	GAAATCACGG	TAGTCCTCTA	TGATGGACCC	GCGTCACCTT	GACAGGACCC
101	ACTACCGGCA	AAGTGAAGTC	CTCCGTGAGC	TGCACGTGGA	CACCAGATTT
	TGATGGCCGT	TTCACCTGAG	GAGGCACTCG	ACGTGCACCT	GTGGTCTAAA
151	CCTGCTACAG	CGCCAGGAGC	TCTTCCGTTG	GGCCCGTCAG	TCCTGTACAA
	GGACGATGTC	GCGGTCCTCG	AGAAGGCAAC	CCGGGCAGTC	AGGACATGTT
201	AAAGACTGTG	TTCGTAGAGT	TCACGGATCA	ACTTTTTCAGC	GTTGCCAGGC
	TTTCTGACAC	AAGCATCTCA	AGTGCCTAGT	TGAAAAGTCG	CAACGGTCCG
251	CCAGGCCACC	ATGGATGGGT	CTGCTGGGTC	CTACCATCCA	GGCTGAGGTT
	GGTCCGGTGG	TACCTACCCA	GACGACCCAG	GATGGTAGGT	CCGACTCCAA
301	TACGACACGG	TGGTCGTTAC	CCTGAAGAAC	ATGGCTTCTC	ATCCCGTTAG
	ATGCTGTGCC	ACCAGCAATG	GGACTTCTTG	TACCGAAGAG	TAGGGCAATC
351	TCTTCACGCT	GTCGGCGTCT	CCTTCTGGAA	ATCTTCCGAA	GGCGCTGAAT
	AGAAGTGCGA	CAGCCGCAGA	GGAAGACCTT	TAGAAGGCTT	CCGCGACTTA
401	ATGAGGATCA	CACCAGCCAA	AGGGAGAAGG	AAGACGATAA	AGTCCTTCCC
	TACTCCTAGT	GTGGTCGGTT	TCCCTCTTCC	TTCTGCTATT	TCAGGAAGGG
451	GGTAAAAGCC	AAACCTACGT	CTGGCAGGTC	CTGAAAGAAA	ATGGTCCAAC
	CCATTTTCGG	TTTGGATGCA	GACCGTCCAG	GACTTTCTTT	TACCAGGTTG
501	AGCCTCTGAC	CCACCATGTC	TTACCTACTC	ATACCTGTCT	CACGTGGACC
	TCGGAGACTG	GGTGGTACAG	AATGGATGAG	TATGGACAGA	GTGCACCTGG
551	TGGTGAAAGA	CCTGAATTCT	GGCCTCATTG	GAGCCCTGCT	GGTTTGTAGA
	ACCACTTTCT	GGACTTAAGC	CCGGAGTAAC	CTCGGGACGA	CCAAACATCT
601	GAAGGGAGTC	TGACCAGAGA	AAGGACCCAG	AACCTGCACG	AATTTGTAAT
	TCTCCCTCAG	ACTGGTCTCT	TTCTGGGTC	TTGGACGTGC	TTAAACATGA
651	ACTTTTTTGCT	GTCTTTGATG	AAGGGAAAAG	TTGGCACTCA	GCAAGAAATG
	TGAAAAACGA	CAGAAACTAC	TTCCCTTTTC	AACCGTGAGT	CGTTCTTTAC
701	ACTCCTGGAC	ACGGGCCATG	GATCCCGCAC	CTGCCAGGGC	CCAGCCTGCA
	TGAGGACCTG	TGCCCCGTAC	CTAGGGCGTG	GACGGTCCCG	GGTCGGACGT
751	ATGCACACAG	TCAATGGCTA	TGTCAACAGG	TCTCTGCCAG	GTCTGATCGG
	TACGTGTGTC	AGTTACCGAT	ACAGTTGTCC	AGAGACGGTC	CAGACTAGCC
801	ATGTCATAAG	AAATCAGTCT	ACTGGCACGT	GATTGGAATG	GGCACCAGCC
	TACAGTATTC	TTTAGTCAGA	TGACCGTGCA	CTAACCTTAC	CCGTGGTCGG
851	CGGAAGTGCA	CTCCATTTTT	CTTGAAGGCC	ACACGTTTCT	CGTGAGGCAC
	GCCTTCACGT	GAGGTAAAAA	GAACCTCCGG	TGTGCAAAGA	GCACTCCGTG
901	CATCGCCAGG	CTTCCTTGGA	GATCTCGCCA	CTAACTTTCC	TCACTGCTCA
	GTAGCGGTCC	GAAGGAACCT	CTAGAGCGGT	GATTGAAAGG	AGTGACGAGT
951	GACATTCCCTG	ATGGACCTTG	GCCAGTTCCT	ACTGTTTTGT	CATATCTCTT
	CTGTAAGGAC	TACCTGGAAC	CGGTCAAGGA	TGACAAAACA	GTATAGAGAA
1001	CCCACCACCA	TGGTGGCATG	GAGGCTCACG	TCAGAGTAGA	AAGCTGCGCC
	GGGTGGTGGT	ACCACCGTAC	CTCCGAGTGC	AGTCTCATCT	TTCGACGCGG
1051	GAGGAGCCCC	AGCTGCGGAG	GAAAGCTGAT	GAAGAGGAAG	ATTATGATGA
	CTCCTCGGGG	TCGACGCCTC	CTTTCGACTA	CTTCTCCTTC	TAATACTACT
1101	CAATTTGTAC	GACTCGGACA	TGGACGTGGT	CCGGCTCGAT	GGTGACGACG
	GTTAAACATG	CTGAGCCTGT	ACCTGCACCA	GGCCGAGCTA	CCACTGCTGC
1151	TGTCTCCCTT	TATCCAAATC	CGCTCGGTTG	CCAAGAAGCA	TCCTAAAACT
	ACAGAGGGAA	ATAGGTTTAG	GCGAGCCAAC	GGTTCTTCGT	AGGATTTTGA
1201	TGGGTACATT	ACATTGCTGC	TGAAGAGGAG	GACTGGGACT	ATGCTCCCTT

FIG. 9A

21/33

1251	ACCCATGTAA	TGTAACGACG	ACTTCTCCTC	CTGACCCTGA	TACGAGGGAA
	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAATG
	TCAGGAGCGG	GGGCTACTGT	CTTCAATATT	TTCAGTTATA	AACTTGTTAC
1301	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAA	AAGTCCGATT	TATGGCATAAC
	CGGGAGTCGC	CTAACCATCC	TTCATGTTTT	TTCAGGCTAA	ATACCGTATG
1351	ACAGATGAAA	CCTTTAAGAC	GCGTGAAGCT	ATTCAGCATG	AATCAGGAAT
	TGTCTACTTT	GGAAATTCTG	CGCACTTCGA	TAAGTCGTAC	TTAGTCCTTA
1401	CTTGGGACCT	TTACTTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
	GAACCCTGGA	AATGAAATAC	CCCTTCAACC	TCTGTGTGAC	AACTAATATA
1451	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	TCTACCCTCA	CGGAATCACT
	AATTCTTAGT	TCGTTCGTCT	GGTATATTGT	AGATGGGAGT	GCCTTAGTGA
1501	GATGTCCGTC	CTTTGTATTC	AAGGAGATTA	CCAAAAGGTG	TAAAACATTT
	CTACAGGCAG	GAAACATAAG	TTCTCTAAT	GGTTTTCCAC	ATTTTGTAAG
1551	GAAGGATTTT	CCAATTCTGC	CAGGAGAAAT	ATTCAAATAT	AAATGGACAG
	CTTCCTAAAA	GGTTAAGACG	GTCCTCTTTA	TAAGTTTATA	TTTACCTGTC
1601	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCGCGGTG	CCTGACCCGC
	ACTGACATCT	TCTACCCGGT	TGATTTAGTC	TAGGCGCCAC	GGACTGGGCG
1651	TATTACTCTA	GTTTCGTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
	ATAATGAGAT	CAAAGCAATT	ATACCTCTCT	CTAGATCGAA	GTCCTGAGTA
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAATC	TGTAGATCAA	AGAGGAAACC
	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTAGTT	TCTCCTTTGG
1751	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTTCTGT	ATTTGATGAG
	TCTATTACAG	TCTGTTCTCC	TTACAGTAGG	ACAAAAGACA	TAAACTACTC
1801	AACCGAAGCT	GGTACCTCAC	AGAGAATATA	CAACGCTTTC	TCCCCAATCC
	TTGGCTTCGA	CCATGGAGTG	TCTCTTATAT	GTTGCGAAAG	AGGGGTTAGG
1851	AGCTGGAGTG	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
	TCGACCTCAC	GTCGAACTCC	TAGGTCTCAA	GGTTCGGAGG	TTGTAGTACG
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTTGTTTG
	TGTCGTAGTT	ACCGATACAA	AAACTATCAA	ACGTCAACAG	TCAAACAAAC
1951	CATGAGGTGG	CATACTGGTA	CATTCTAAGC	ATTGGAGCAC	AGACTGACTT
	GTACTCCACC	GTATGACCAT	GTAAGATTCC	TAACCTCGTG	TCTGACTGAA
2001	CCTTTCTGTC	TTCTTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
	GGAAAGACAG	AAGAAGAGAC	CTATATGGAA	GTTTGTGTTT	TACCAGATAC
2051	AAGACACACT	CACCCATTTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTGC
	TTCTGTGTGA	GTGGGATAAG	GGTAAGAGTC	CTCTTTGACA	GAAGTACAGC
2101	ATGGAAGAAC	CAGGTCTATG	GATTCTGGGG	TGCCACAAC	CAGACTTTTCG
	TACCTTTTGG	GTCCAGATAC	CTAAGACCCC	ACGGTGTGTA	GTCTGAAAGC
2151	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAACA
	CTTGTCTCCG	TACTGGCGGA	ATGACTTCCA	AAGATCAACA	CTGTTCTTGT
2201	CTGGTGATTA	TTACGAGGAC	AGTTATGAAG	ATATTTTCAGC	ATACTTGCTG
	GACCACTAAT	TTACGAGGAC	TCAATACTTC	TATAAAGTCG	TATGAACGAC
2251	AGTAAAAACA	ATGCCATTGA	ACCTAGGAGC	TTTGCCCGAG	ATTCAAGACC
	TCATTTTTGT	TACGGTAACT	TGGATCCTCG	AAACGGGTCT	TAAGTTCTGG
2301	CCCTAGTGCG	AGCGCTCCAA	AGCCTCCGGT	CCTGCGACGG	CATCAGAGGG
	GGGATCACGC	TCGCGAGGTT	TCGGAGGCCA	GGACGCTGCC	GTAGTCTCCC
2351	ACATAAGCCT	TCCTACTTTT	CAGCCGGAGG	AAGACAAAAT	GGACTATGAT
	TGTATTTCGA	AGGATGAAAA	GTCGGCCTCC	TTCTGTTTTA	CCTGATACTA
2401	GATATCTTCT	CAACTGAAAC	GAAGGGAGAA	GATTTTGACA	TTTACGGTGA
	CTATAGAAGA	GTTGACTTTG	CTTCCCTCTT	CTAAAAC	AAATGCCACT
2451	GGATGAAAAT	CAGGACCCTC	GCAGCTTTCA	GAAGAGAACC	CGACACTATT
	CCTACTTTTA	GTCCTGGGAG	CGTCGAAAGT	CTTCTCTTGG	GCTGTGATAA
2501	TCATTGCTGC	GGTGGAGCAG	CTCTGGGATT	ACGGGATGAG	CGAATCCCCC

FIG. 9B

22/33

2551	AGTAACGACG	CCACCTCGTC	GAGACCCTAA	TGCCCTACTC	GCTTAGGGGG
	CGGGCGCTAA	GAAACAGGGC	TCAGAACGGA	GAGGTGCCTC	GGTTCAAGAA
	GCCCCGCGATT	CTTTGTCCCC	AGTCTTGCCCT	CTCCACGGAG	CCAAGTTCTT
2601	GGTGGTCTTC	CGGGAATTTG	CTGACGGCTC	CTTCACGCAG	CCGTTCGTACC
	CCACCAGAAG	GCCCTTAAAC	GACTGCCGAG	GAAGTGCCTC	GGCAGCATGG
2651	GCGGGGAACT	CAACAAACAC	TTGGGGCTCT	TGGGACCCTA	CATCAGAGCG
	CGCCCCCTTGA	GTTGTTTGTG	AACCCCGAGA	ACCCTGGGAT	GTAGTCTCGC
2701	GAAGTTGAAG	ACAACATCAT	GGTAACTTTC	AAAAACCAGG	CGTCTCGTCC
	CTTCAACTTC	TGTTGTAGTA	CCATTGAAAG	TTTTTGGTCC	GCAGAGCAGG
2751	CTATTCCTTC	TACTCGAGCC	TTATTTCTTA	TCCGGATGAT	CAGGAGCAAG
	GATAAGGAAG	ATGAGCTCGG	AATAAAGAAT	AGGCCTACTA	GTCCTCGTTC
2801	GGGCAGAACC	TCGACACAAC	TTCGTCCAGC	CAAATGAAAC	CAGAACTTAC
	CCCGTCTTGG	AGCTGTGTTG	AAGCAGGTCG	GTTTACTTTG	GTCTTGAATG
2851	TTTTGGAAAG	TGCAGCATCA	CATGGCACCC	ACAGAAGACG	AGTTTGACTG
	AAAACCTTTC	ACGTCGTAGT	GTACCGTGGG	TGTCTTCTGC	TCAAACGTAC
2901	CAAAGCCTGG	GCCTACTTTT	CTGATGTTGA	CCTGGAAAAA	GATGTGCACT
	GTTTCGGACC	CGGATGAAAA	GACTACAAC	GGACCTTTTT	CTACACGTGA
2951	CAGGCTTGAT	CGGCCCCCTT	CTGATCTGCC	GCGCCAACAC	CCTGAACGCT
	GTCCGAAC	GCCGGGGGAA	GACTAGACGG	CGCGGTGTG	GGACTTGCGA
3001	GCTCACGGTA	GACAAGTGAC	CGTGCAAGAA	TTTGCTCTGT	TTTTCACTAT
	CGAGTGCCAT	CTGTTCACTG	GCACGTTCTT	AAACGAGACA	AAAAGTGATA
3051	TTTTGATGAG	ACAAAGAGCT	GGTACTTCAC	TGAAAAATGTG	GAAAGGAACT
	AAAAC	TGTTTCTCGA	CCATGAAGTG	ACTTTTACAC	CTTTCCTTGA
3101	GCCGGGCCCC	CTGCCATCTG	CAGATGGAGG	ACCCCACTCT	GAAAGAAAAAC
	CGGCCCCGGG	GACGGTAGAC	GTCTACCTCC	TGGGGTGAGA	CTTTCCTTTTG
3151	TATCGCTTCC	ATGCAATCAA	TGGCTATGTG	ATGGATACAC	TCCCTGGCTT
	ATAGCGAAGG	TACGTTAGTT	ACCGATACAC	TACCTATGTG	AGGGACCGAA
3201	AGTAATGGCT	CAGAATCAAA	GGATCCGATG	GTATCTGCTC	AGCATGGGCA
	TCATTACCGA	GTCTTAGTTT	CCTAGGCTAC	CATAGACGAG	TCGTACCCGT
3251	GCAATGAAAA	TATCCATTCTG	ATTCAATTTTA	GCGGACACGT	GTTTCACTGTA
	CGTTACTTTT	ATAGGTAAGC	TAAGTAAAT	CGCCTGTGCA	CAAGTCACAT
3301	CGGAAAAAGG	AGGAGTATAA	AATGGCCGTG	TACAATCTCT	ATCCGGGTGT
	GCCTTTTTC	TCCTCATATT	TTACCGGCAC	ATGTTAGAGA	TAGGCCCA
3351	CTTTGAGACA	GTGGAAATGC	TACCGTCCAA	AGTTGGAATT	TGGCGAATAG
	GAAACTCTGT	CACCTTTACG	ATGGCAGGTT	TCAACCTTAA	ACCGCTTATC
3401	AATGCCTGAT	TGGCGAGCAC	CTGCAAGCTG	GGATGAGCAC	GACTTTCCTG
	TTACGGACTA	ACCGCTCGTG	GACGTTTCGAC	CCTACTCGTG	CTGAAAGGAC
3451	GTGTACAGCA	AGAAGTGTC	GACTCCCCTG	GGAATGGCTT	CTGGACACAT
	CACATGTCGT	TCTTCACAGT	CTGAGGGGAC	CCTTACCGAA	GACCTGTGTA
3501	TAGAGATTTT	CAGATTACAG	CTTCAGGACA	ATATGGACAG	TGGGCCCA
	ATCTCTAAAA	GTCTAATGTC	GAAGTCCTGT	TATACCTGTC	ACCCGGGGTT
3551	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA	TCAATGCCTG	GAGCACCAAG
	TCGACCGGTC	TGAAGTAATA	AGGCCTAGTT	AGTTACGGAC	CTCGTGGTTC
3601	GAGCCCTTTT	CTTGGATCAA	GGTGGATCTG	TTGGCACCAA	TGATTATTCA
	CTCGGGAAAA	GAACCTAGTT	CCACCTAGAC	AACCGTGGTT	ACTAATAAGT
3651	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA	GTTCTCCAGC	CTCTACATCT
	GCCGTAGTTC	TGGGTCCAC	GGGCAGTCTT	CAAGAGGTCG	GAGATGTAGA
3701	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG	GGAAGAAGTG	GCAGACTTAT
	GAGTCAAATA	GTAGTACATA	TCAGAACTAC	CCTTCTTCAC	CGTCTGAATA
3751	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC	TTCTTTGGCA	ATGTGGATTTC
	GCTCCTTTAA	GGTGACCTTG	GAATTACCAG	AAGAAACCGT	TACACCTAAG
3801	ATCTGGGATA	AAACACAATA	TTTTTAACCC	TCCAATTATT	GCTCGATACA

FIG. 9C

23/33

	TAGACCCTAT	TTTGTGTTAT	AAAAATTGGG	AGGTTAATAA	CGAGCTATGT
3851	TCCGTTTGCA	CCCAACTCAT	TATAGCATTC	GCAGCACTCT	TCGCATGGAG
	AGGCAAACGT	GGGTTGAGTA	ATATCGTAAG	CGTCGTGAGA	AGCGTACCTC
3901	TTGATGGGCT	GTGATTTAAA	TAGTTGCAGC	ATGCCATTGG	GAATGGAGAG
	AACTACCCGA	CACTAAATTT	ATCAACGTCG	TACGGTAACC	CTTACCTCTC
3951	TAAAGCAATA	TCAGATGCAC	AGATTACTGC	TTCATCCTAC	TTTACCAATA
	ATTTTCGTTAT	AGTCTACGTG	TCTAATGACG	AAGTAGGATG	AAATGGTTAT
4001	TGTTTGCCAC	CTGGTCTCCT	TCAAAAGCTC	GACTTCACCT	CCAAGGGAGG
	ACAAACGGTG	GACCAGAGGA	AGTTTTTCGAG	CTGAAGTGGA	GGTTCCCTCC
4051	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT	CCAAAAGAGT	GGCTGCAAGT
	TCATTACGGA	CCTCTGGAGT	CCACTTATTA	GGTTTTCTCA	CCGACGTTCA
4101	GGACTTCCAG	AAGACAATGA	AAGTCACAGG	AGTAACTACT	CAGGGAGTAA
	CCTGAAGGTC	TTCTGTACT	TTCAGTGTCC	TCATTGATGA	GTCCCTCATT
4151	AATCTCTGCT	TACCAGCATG	TATGTGAAGG	AGTTCCTCAT	CTCCAGCAGT
	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC	TCAAGGAGTA	GAGGTCGTCA
4201	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTTT	CAGAATGGCA	AAGTAAAGGT
	GTTCTACCGG	TAGTCACCTG	AGAGAAAAAA	GTCTTACCGT	TTCATTTCCA
4251	TTTTTCAGGA	AATCAAGACT	CCTTCACACC	TGTGGTGAAC	TCTCTAGACC
	AAAAGTCCCT	TTAGTTCTGA	GGAAGTGTGG	ACACCACTTG	AGAGATCTGG
4301	CACCGTTACT	GACTCGCTAC	CTTCGAATTC	ACCCCCAGAG	TTGGGTGCAC
	GTGGCAATGA	CTGAGCGATG	GAAGCTTAAG	TGGGGGTCTC	AACCCACGTG
4351	CAGATTGCCC	TGAGGATGGA	GGTTCTGGGC	TGCGAGGCAC	AGGACCTCTA
	GTCTAACGGG	ACTCCTACCT	CCAAGACCCG	ACGCTCCGTG	TCCTGGAGAT
4401	C				
	G				

FIG. 9D

24/33

AMINO ACID SEQUENCE OF HUMAN B DOMAIN-DELETED FACTOR VIII (HSQ)

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe
1				5					10					15	
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser
			20					25					30		
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg
		35					40					45			
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val
	50					55					60				
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Val	His	Leu	Phe	Asn	Ile
65					70					75				80	
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln
				85					90					95	
Ala	Glu	Val	Tyr	Asp	Tyr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
			100					105					110		
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
		115					120					125			
Glu	Gly	Ala	Glu	Thr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
	130					135					140				
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
145					150					155				160	
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
			165						170					175	
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
			180					185					190		
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
		195					200					205			
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
	210					215					220				
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
225					230					235				240	
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
			245						250					255	
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
			260					265					270		
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
		275					280					285			
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
	290					295					300				
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
305					310					315					320
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
				325					330					335	
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
			340					345					350		
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
		355					360					365			
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
	370					375					380				
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr
385					390					395					400
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro

FIG. 10A

25/33

Leu	Val	Leu	Ala	405	Pro	Asp	Asp	Arg	Ser	410	Tyr	Lys	Ser	Gln	Tyr	415	Leu	Asn
			420						425						430			
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met			
		435					440						445					
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu			
	450					455					460							
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu			
	465				470					475					480			
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro			
			485						490					495				
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys			
			500					505					510					
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe			
		515					520					525						
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp			
	530					535					540							
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg			
	545				550					555				560				
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu			
			565					570						575				
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val			
			580					585					590					
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu			
		595					600					605						
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp			
	610					615					620							
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val			
	625				630				635					640				
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp			
			645					650					655					
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe			
			660					665					670					
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr			
		675					680					685						
Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro			
	690					695					700							
Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly			
	705				710					715				720				
Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp			
			725					730					735					
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys			
			740					745					750					
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu			
		755					760					765						
Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln			
	770					775					780							
Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Lys	Glu			
	785				790				795					800				
Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe			
			805					810					815					
Gln	Lys	Lys	Thr	Arg	His	Hyr	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp			
			820					825				830						
Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Atg	Ala	Gln			
		835					840					845						
Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr			
	850					855					860							

FIG. 10B

26/33

Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu	His
865					870					875					880
Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	Arg	Ala	Glu	Val	Glu	Asp	Asn	Ile
				885					890					895	
Met	Val	Thr	Phe	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser
			900					905					910		
Ser	Leu	Ile	Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg
		915					920					925			
Lys	Asn	Phe	Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val
	930					935					940				
Gln	His	His	Met	Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp
945					950					955					960
Ala	Tyr	Phe	Ser	Asp	Val	Asp	Leu	Glu	Lys	Asp	Val	His	Ser	Gly	Leu
				965					970					975	
Ile	Gly	Pro	Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	Ala	His
			980					985					990		
Gly	Arg	Gln	Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	Thr	Ile	Phe
		995					1000					1005			
Asp	Glu	Thr	Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	Glu	Arg	Asn	Cys
	1010					1015					1020				
Arg	Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	Lys	Glu	Asn
1025					1030					1035					1040
Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	Thr	Leu	Pro	Gly
				1045					1050					1055	
Leu	Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	Tyr	Leu	Leu	Ser	Met
			1060					1065					1070		
Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	Phe	Ser	Gly	His	Val	Phe
		1075					1080					1085			
Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	Met	Ala	Leu	Tyr	Asn	Leu	Tyr
	1090					1095						1100			
Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	Met	Leu	Pro	Ser	Lys	Ala	Gly	Ile
1105					1110					1115					1120
Trp	Arg	Val	Glu	Cys	Leu	Ile	Gly	Glu	His	Leu	His	Ala	Gly	Met	Ser
				1125					1130					1135	
Thr	Leu	Phe	Leu	Val	Tyr	Ser	Asn	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Met
			1140					1145					1150		
Ala	Ser	Gly	His	Ile	Arg	Asp	Phe	Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr
		1155					1160					1165			
Gly	Gln	Trp	Ala	Pro	Lys	Leu	Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser	Ile
	1170					1175					1180				
Asn	Ala	Trp	Ser	Thr	Lys	Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu
1185					1190					1195					1200
Leu	Ala	Pro	Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln
				1205					1210					1215	
Lys	Phe	Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu
			1220					1225					1230		
Asp	Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	Leu
		1235					1240					1245			
Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	Asn	Ile
	1250					1255					1260				
Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	Pro	Thr	His
1265					1270					1275					1280
Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met	Gly	Cys	Asp	Leu
				1285					1290					1295	
Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser	Lys	Ala	Ile	Ser	Asp
			1300					1305					1310		
Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr	Asn	Met	Phe	Ala	Thr	Trp

FIG. 10C

28/33

NUCLEOTIDE SEQUENCE OF HUMAN B DOMAIN-DELETED FACTOR VIII(HSQ)

```
1  ATGCAAATAG AGCTCTCCAC CTGCTTCTTT CTGTGCCTTT TGCGATTCTG
51  CTTTAGTGCC ACCAGAAGAT ACTACCTGGG TGCAGTGGAA CTGTCATGGG
101 ACTATATGCA AAGTGATCTC GGTGAGCTGC CTGTGGACGC AAGATTTTCT
151 CCTAGAGTGC CAAAATCTTT TCCATTCAAC ACCTCAGTCG TGTACAAAAA
201 GACTCTGTTT GTAGAATTCA CGGTTACACT TTTCAACATC GCTAAGCCAA
251 GGCCACCCTG GATGGGTCTG CTAGGTCCTA CCATCCAGGC TGAGGTTTAT
301 GATACAGTGG TCATTACACT TAAGAACATG GCTTCCCATC CTGTCAGTCT
351 TCATGCTGTT GGTGTATCCT ACTGGAAAGC TTCTGAGGGA GCTGAATATG
401 ATGATCAGAC CAGTCAAAGG GAGAAAGAAG ATGATAAAGT CTTCCCTGGT
451 GGAAGCCATA CATATGTCTG GCAGGTCTCT AAAGAGAATG GTCCAATGGC
501 CTCTGACCCA CTGTGCCTTA CCTACTCATA TCTTTCTCAT GTGGACCTGG
551 TAAAAGACTT GAATTCAGGC CTCATTGGAG CCCTACTAGT ATGTAGAGAA
601 GGGAGTCTGG CCAAGGAAAA GACACAGACC TTGCACAAAT TTATACTACT
651 TTTTGCTGTA TTTGATGAAG GGAAAAGTTG GCACTCAGAA ACAAAGAACT
701 CCTTGATGCA GGATAGGGAT GCTGCATCTG CTCGGGCTCT GCCTAAAATG
751 CACACAGTCA ATGTTTATGT AAACAGGTCT CTGCCAGTCT TGATTGGATG
801 CCACAGGAAA TCAGTCTATT GGCATGTGAT TGGAAATGGG ACCACTCCTG
851 AAGTGCCTC AATATTCCTC GAAGGTCACA CATTTCTTGT GAGGAACCAT
901 CGCCAGGCGT CCTTGGAAAT CTCGCCAATA ACTTTCCTTA CTGCTCAAAC
951 ACTCTTGATG GACCTTGGAC AGTTTCTACT GTTTTGTCAT ATCTCTTCCC
1001 ACCAACATGA TGGCATGGAA GCTTATGTCA AAGTAGACAG CTGTCCAGAG
1051 GAACCCCAAC TACGAATGAA AAATAATGAA GAAGCGGAAG ACTATGATGA
1101 TGATCTTACT GATTCTGAAA TGGATGTGGT CAGGTTTGAT GATGACAACT
1151 CTCCTTCCTT TATCCAAATT CGCTCAGTTG CCAAGAAGCA TCCTAAAAC
1201 TGGGTACATT ACATTGCTGC TGAAGAGGAG GACTGGGACT ATGCTCCCTT
1251 AGTCCTCGCC CCCGATGACA GAAGTTATAA AAGTCAATAT TTGAACAATG
1301 GCCCTCAGCG GATTGGTAGG AAGTACAAAA AAGTCCGATT TATGGCATA
1351 ACAGATGAAA CCTTTAAGAC GCGTGAAGCT ATTCAGCATG AATCAGGAAT
1401 CTTGGGACCT TTACTTTATG GGGAAAGTTG AGACACACTG TTGATTATAT
1451 TTAAGAATCA AGCAAGCAGA CCATATAACA TCTACCCTCA CGGAATCACT
1501 GATGTCCGTC CTTTGTATTC AAGGAGATTA CCAAAGGGTG TAAAACATTT
1551 GAAGGATTTT CCAATTCTGC CAGGAGAAAT ATTCAAATAT AAATGGACAG
1601 TGACTGTAGA AGATGGGCCA ACTAAATCAG ATCCGCGGTG CCTGACCCGC
1651 TATTACTCTA GTTTCGTAA TATGGAGAGA GATCTAGCTT CAGGACTCAT
1701 TGGCCCTCTC CTCATCTGCT ACAAAGAATC TGTAGATCAA AGAGGAAACC
1751 AGATAATGTC AGACAAGAGG AATGTCATCC TGTTTTCTGT ATTTGATGAG
1801 AACCGAAGCT GGTACCTCAC AGAGAATATA CAACGCTTTC TCCCCAATCC
1851 AGCTGGAGTG CAGCTTGAGG ATCCAGAGTT CCAAGCCTCC AACATCATGC
1901 ACAGCATCAA TGGCTATGTT TTTGATAGTT TGCAGTTGTC AGTTTGTTTG
1951 CATGAGGTGG CATACTGGTA CATTCTAAGC ATTGGAGCAC AGACTGACTT
2001 CCTTTCTGTC TTCTTCTCTG GATATACTTT CAAACACAAA ATGGTCTATG
2051 AAGACACACT CACCCTATTC CCATTCTCAG GAGAAACTGT CTTCTATGTC
2100 ATGGAAAACC CAGGTCTATG GATTCTGGGG TGCCACAACCT CAGACTTTTC
2151 GAACAGAGGC ATGACCGCCT TACTGAAGGT TTCTAGTTGT GACAAGAACA
2201 CTGGTGATTA TTACGAGGAC AGTTATGAAG ATATTTTCAGC ATACTTGCTG
2251 AGTAAAAACA ATGCCATTGA ACCTAGGAGC TTCTCTCAGA ATCCACCAGT
2301 CTTGAAACGC CATCAACGGG AAATAACTCG TACTACTCTT CAGTCAGATC
2351 AAGAGGAAAT TGACTATGAT GATACCATAT CAGTTGAAAT GAAGAAGGAA
2401 GATTTTGACA TTTATGATGA GGATGAAAAT CAGAGCCCCC GCAGCTTTCA
```

FIG. 11A

29/33

2451	AAAGAAAACA	CGACACTATT	TTATTGCTGC	AGTGGAGAGG	CTCTGGGATT
2501	ATGGGATGAG	TAGCTCCCCA	CATGTTCTAA	GAAACAGGGC	TCAGAGTGGC
2551	AGTGTCCCTC	AGTTCAAGAA	AGTTGTTTTT	CAGGAATTTA	CTGATGGCTC
2601	CTTTACTCAG	CCCTTATACC	GTGGAGAACT	AAATGAACAT	TTGGGACTCC
2651	TGGGGCCATA	TATAAGAGCA	GAAGTTGAAG	ATAATATCAT	GGTAACTTTC
2701	AGAAATCAGG	CCTCTCGTCC	CTATTCCTTC	TATTCTAGCC	TTATTTCTTA
2751	TGAGGAAGAT	CAGAGGCAAG	GAGCAGAACC	TAGAAAAAAC	TTTGTCAAGC
2801	CTAATGAAAC	CAAAACTTAC	TTTTGGAAAG	TGCAACATCA	TATGGCACCC
2851	ACTAAAGATG	AGTTTGACTG	CAAAGCCTGG	GCTTATTTCT	CTGATGTTGA
2901	CCTGGAAAAA	GATGTGCACT	CAGGCCTGAT	TGGACCCCTT	CTGGTCTGCC
2951	ACACTAACAC	ACTGAACCCT	GCTCATGGGA	GACAAGTGAC	AGTACAGGAA
3001	TTTGCTCTGT	TTTTCACCAT	CTTTGATGAG	ACCAAAAGCT	GGTACTTCAC
3051	TGAAAATATG	GAAAGAAACT	GCAGGGCTCC	CTGCAATATC	CAGATGGAAG
3101	ATCCCACCTT	TAAAGAGAAT	TATCGCTTCC	ATGCAATCAA	TGGCTACATA
3151	ATGGATACAC	TACCTGGCTT	AGTAATGGCT	CAGGATCAAA	GGATTGATG
3201	GTATCTGCTC	AGCATGGGCA	GCAATGAAAA	CATCCATTCT	ATTCATTTCA
3251	GTGGACATGT	GTTCACTGTA	CGAAAAAAAG	AGGAGTATAA	AATGGCACTG
3301	TACAATCTCT	ATCCAGGTGT	TTTTGAGACA	GTGGAAATGT	TACCATCCAA
3351	AGCTGGAATT	TGGCGGGTGG	AATGCCTTAT	TGGCGAGCAT	CTACATGCTG
3401	GGATGGACAC	ACTTTTTCTG	GTGTACAGCA	ATAAGTGTC	GACTCCCTG
3451	GGAATGGCTT	CTGGACACAT	TAGAGATTTT	CAGATTACAG	CTTCAGGACA
3501	ATATGGACAG	TGGGCCCCAA	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA
3551	TCAATGCCTG	GAGCACCAAG	GAGCCCTTTT	CTTGGATCAA	GGTGGATCTG
3601	TTGGCACCAA	TGATTATTCA	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA
3651	GTTCTCCAGC	CTCTACATCT	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG
3701	GGAAGAAGTG	GCAGACTTAT	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC
3751	TTCTTTGGCA	ATGTGGATTTC	ATCTGGGATA	AAACACAATA	TTTTTAACCC
3801	TCCAATTATT	GCTCGATACA	TCCGTTTGCA	CCCAACTCAT	TATAGCATTC
3851	GCAGCACTCT	TCGCATGGAT	TTGATGGGCT	GTGATTTAAA	TAGTTGCAGC
3901	ATGCCATTGG	GAATGGAGAG	TAAAGCAATA	TCAGATGCAC	AGATTACTGC
3951	TTCATCCTAC	TTTACCAATA	TGTTTGCCAC	CTGGTCTCCT	TCAAAAGCTC
4001	GACTTCACCT	CCAAGGGAGG	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT
4051	CCAAAAGAGT	GGCTGCAAGT	GGACTTCCAG	AAGACAATGA	AAGTCACAGG
4101	AGTAACTACT	CAGGGAGTAA	AATCTCTGCT	TACCAGCATG	TATGTGAAGG
4151	AGTTCCTCAT	CTCCAGCAGT	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTTT
4201	CAGAATGGCA	AAGTAAAGGT	TTTTCAGGGA	AATCAAGACT	CCTTCACACC
4251	TGTGGTGAAC	TCTCTAGACC	CACCGTTACT	GACTCGCTAC	CTTCGAATTC
4301	ACCCCAGAG	TTGGGTGCAC	CAGATTGCCC	TGAGGATGGA	GGTTCTGGGC
4351	TGCGAGGCAC	AGGACCTCTA	C		

FIG. 11B

30/33

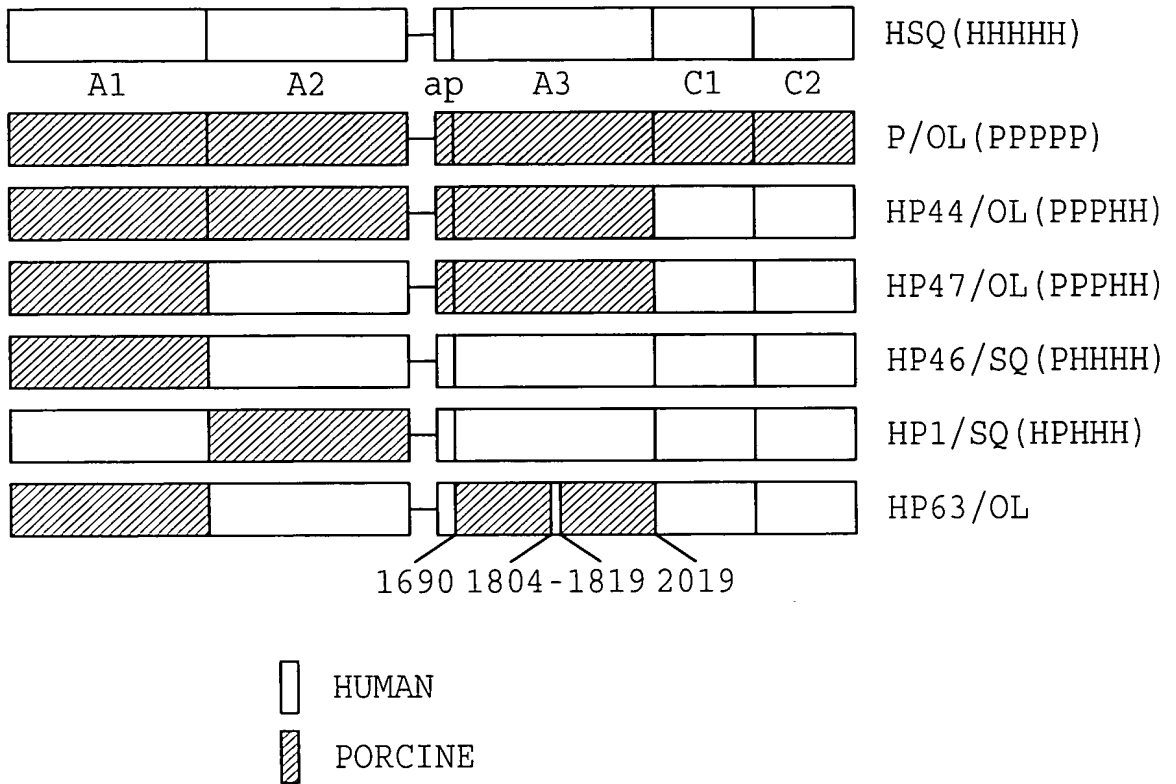


FIG. 12

AMINO ACID SEQUENCE OF HP630L

1	MQLELSTCVF	LCLLPGLGFS	IRRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKT	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSS	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYVWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLFC	HISHHHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRLD	GDDVSPFIQI	RSVAKKHPKT
401	WVHYIAAEEE	DWDYAPLVLA	PDDRSYKSOY	LNNGPQRIGR	KYKKVRFMAY
451	TDETFKTR	IQHESGILGP	LLYGEVGD	LIIFKNQASR	PYNIYPHGIT
501	DVRPLYSRRL	PKGVKHLKDF	PILPGEIFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSFVNMER	DLASGLIGPL	LICYKESVDQ	RGNQIMSDKR	NVILFSVFDE
601	NRSWYLTE	QRFLPNPAGV	QLEDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYILS	IGAQTDFLSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWILG	CHNSDFRNRG	MTALLKVSSC	DKNTGDYYED	SYEDISAYLL
751	SKNNAIEPRS	FSQNSRHPST	RSQNPPVLKR	HQREITRTTL	QSDQEEIDYD
801	DTISVEMKKE	DFDIYDEDEN	QSPRSFQKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALRNRAQNG	EVPRFKKVVF	REFADGSFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVTF	KNQASRPYSF	YSSLISYPDD	QEQGAEPKRN	FVKPNETKTY
951	FWKVQHMAP	TEDEFDCKAW	AYFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLKEN
1051	YRFHAINGYV	MDTLPGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFSV
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVG	WRNRCLIGEH	LQAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASGOYGO	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKVDL	LAPMIIHG	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNIFNPPII	ARYIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLQGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISSS
1401	QDGHQWTLFF	QNGKVKVFGQ	NQDSFTPVVN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY			

FIG. 13

32/33

NUCLEOTIDE SEQUENCE OF HP63/OL

```
1  ATGCAGCTAG AGCTCTCCAC CTGTGTCTTT CTGTGTCTCT TGCCACTCGG
51  CTTTAGTGCC ATCAGGAGAT ACTACCTGGG CGCAGTGGA CTGTCTCTGGG
101 ACTACCGGCA AAGTGAAGTC CTCCGTGAGC TGCACGTGGA CACCAGATTT
151 CCTGCTACAG CGCCAGGAGC TCTTCCGTTG GGCCCGTCAG TCCTGTACAA
201 AAAGACTGTG TTCGTAGAGT TCACGGATCA ACTTTTCAGC GTTGCCAGGC
251 CCAGGCCACC ATGGATGGGT CTGCTGGGTC CTACCATCCA GGCTGAGGTT
301 TACGACACGG TGGTCGTTAC CCTGAAGAAC ATGGCTTCTC ATCCCGTTAG
351 TCTTCACGCT GTCGGCGTCT CCTTCTGGAA ATCTTCCGAA GGCGCTGAAT
401 ATGAGGATCA CACCAGCCAA AGGGAGAAGG AAGACGATAA AGTCCTTCCC
451 GGTAAGGCC AAACCTACGT CTGGCAGGTC CTGAAAGAAA ATGGTCCAAC
501 AGCCTCTGAC CCACCATGTC TTACCTACTC ATACCTGTCT CACGTGGACC
551 TGGTGAAAGA CCTGAATTCG GGCCTCATTC GAGCCCTGCT GGTTTGTAGA
601 GAAGGGAGTC TGACCAGAGA AAGGACCCAG AACCTGCACG AATTTGTACT
651 ACTTTTTGCT GTCTTTGATG AAGGGAAAAG TTGGCACTCA GCAAGAAATG
701 ACTCCTGGAC ACGGGCCATG GATCCCGCAC CTGCCAGGGC CCAGCCTGCA
751 ATGCACACAG TCAATGGCTA TGTCAACAGG TCTCTGCCAG GTCTGATCGG
801 ATGTCATAAG AAATCAGTCT ACTGGCACGT GATTGGAATG GGCACCAGCC
851 CGGAAGTGCA CTCCATTTTT CTTGAAGGCC ACACGTTTCT CGTGAGGCAC
901 CATCGCCAGG CTTCTTGGGA GATCTCGCCA CTAACCTTCC TCACTGCTCA
951 GACATTCTTG ATGGACCTTG GCCAGTTCCT ACTGTTTTGT CATATCTCTT
1001 CCCACCACCA TGGTGGCATG GAGGCTCACG TCAGAGTAGA AAGCTGCGCC
1051 GAGGAGCCCC AGCTGCGGAG GAAAGCTGAT GAAGAGGAAG ATTATGATGA
1101 CAATTTGTAC GACTCGGACA TGGACGTGGT CCGGCTCGAT GGTGACGACG
1151 TGTCTCCCTT TATCCAAATC CGCTCAGTTG CCAAGAAGCA TCCTAAAAC
1201 TGGGTACATT ACATTGCTGC TGAAGAGGAG GACTGGGACT ATGCTCCCTT
1251 AGTCCTCGCC CCCGATGACA GAAGTTATAA AAGTCAATAT TTGAACAATG
1301 GCCCTCAGCG GATTGGTAGG AAGTACAAAA AAGTCCGATT TATGGCATA
1351 ACAGATGAAA CCTTTAAGAC TCGTGAAGCT ATTCAGCATG AATCAGGAAT
1401 CTTGGGACCT TTACTTTATG GGAAGTTGG AGACACACTG TTGATTATAT
1451 TTAAGAATCA AGCAAGCAGA CCATATAACA TCTACCCTCA CGGAATCACT
1501 GATGTCCGTC CTTTGTATTC AAGGAGATTA CCAAAGGTG TAAACATTT
1551 GAAGGATTTT CCAATTCTGC CAGGAGAAAT ATTCAAATAT AAATGGACAG
1601 TGAAGTGTAG AGATGGGCCA ACTAAATCAG ATCCTCGGTG CCTGACCCGC
1651 TATTACTCTA GTTTCGTTAA TATGGAGAGA GATCTAGCTT CAGGACTCAT
1701 TGGCCCTCTC CTCATCTGCT ACAAAGAATC TGTAAGTCAA AGAGGAAACC
1751 AGATAATGTC AGACAAGAGG AATGTCATCC TGTTTTCTGT ATTTGATGAG
1801 AACCGAAGCT GGTACCTCAC AGAGAATATA CAACGCTTTC TCCCAATCC
1851 AGCTGGAGTG CAGCTTGAGG ATCCAGAGTT CCAAGCCTCC AACATCATGC
1901 ACAGCATCAA TGGCTATGTT TTTGATAGTT TGCAGTTGTC AGTTTGTGTT
1951 CATGAGGTGG CATACTGGTA CATTCTAAGC ATTGGAGCAC AGACTGACTT
2001 CCTTCTCTGT TTCTTCTCTG GATATACCTT CAAACACAAA ATGGTCTATG
2051 AAGACACACT CACCCTATTC CCATTCTCAG GAGAACTGT CTTCATGTCG
2101 ATGGAAAACC CAGGTCTATG GATTCTGGGG TGCCACAAC CTGACTTTTCG
2151 GAACAGAGGC ATGACCGCCT TACTGAAGGT TTCTAGTTGT GACAAGAACA
2201 CTGGTGATTA TTACGAGGAC AGTTATGAAG ATATTTTCAG ATACTTGCTG
2251 AGTAAAAACA ATGCCATTGA ACCTAGGAGC TTCTCCAGCA ATTCAAGACA
2301 CCCTAGCACT AGGTCTCAAA ACCCACCAGT CTTGAAACGC CATCAACGGG
2351 AAATAACTCG TACTACTCTT CAGTCAGATC AAGAGGAAAT TGACTATGAT
```

FIG. 14A

33/33

```
2401 GATACCATAT CAGTTGAAAT GAAGAAGGAA GATTTTGACA TTTATGATGA
2451 GGATGAAAAT CAGAGCCCCC GCAGCTTTCA AAAGAGAACC CGACACTATT
2501 TCATTGCTGC GGTGGAGCAG CTCTGGGATT ACGGGATGAG CGAATCCCCC
2551 CGGGCGCTAA GAAACAGGGC TCAGAACGGA GAGGTGCCTC GGTTCAAGAA
2601 GGTGGTCTTC CGGGAATTTG CTGACGGCTC CTTACGCAG CCGTCGTACC
2651 GCGGGGAACT CAACAAACAC TTGGGGCTCT TGGGACCTA CATCAGAGCG
2701 GAAGTTGAAG ACAACATCAT GGTAAC TTTC AAAAACCAGG CGTCTCGTCC
2751 CTATTCCTTC TACTCGAGCC TTATTTCTTA TCCGGATGAT CAGGAGCAAG
2801 GGGCAGAACC TCGAAAAAAC TTTGTCAAGC CTAATGAAAC CAAAAC TTAC
2851 TTTTGGAAC TGCAGCATCA CATGGCACCC ACAGAAGACG AGTTTGACTG
2901 CAAAGCCTGG GCCTACTTTT CTGATGTTGA CCTGGAAAAA GATGTGCACT
2951 CAGGCTTGAT CGGCCCCCTT CTGATCTGCC GCGCCAACAC CCTGAACGCT
3001 GCTCACGGTA GACAAGTGAC CGTGCAAGAA TTTGCTCTGT TTTTCACTAT
3051 TTTTGATGAG ACAAAGAGCT GGTACTTCAC TGAAAATGTG GAAAGGAACT
3101 GCCGGGCCCC CTGCCATCTG CAGATGGAGG ACCCCTCT GAAAGAAAAA
3151 TATCGCTTCC ATGCAATCAA TGGCTATGTG ATGGATAAC TCCCTGGCTT
3201 AGTAATGGCT CAGAATCAAA GGATCCGATG GTATCTGCTC AGCATGGGCA
3251 GCAATGAAAA TATCCATTCC ATTCAATTTA GCGGACACGT GTTCAGTGTA
3301 CGGAAAAAGG AGGAGTATAA AATGGCCGTG TACAATCTCT ATCCGGGTGT
3351 CTTTGAGACA GTGGAAATGC TACCGTCCAA AGTTGGAATT TGGCGGAATA
3401 GATGCCTGAT TGGCGAGCAC CTGCAAGCTG GGATGAGCAC GACTTTCCTG
3451 GTGTACAGCA AGAAGTGTC GACTCCCCTG GGAATGGCTT CTGGACACAT
3501 TAGAGATTTT CAGATTACAG CTTCAGGACA ATATGGACAG TGGGCCCCAA
3551 AGCTGGCCAG ACTTCATTAT TCCGGATCAA TCAATGCCTG GAGCACCAG
3601 GAGCCCTTTT CTTGGATCAA GGTGGATCTG TTGGCACCAA TGATTATTCA
3651 CGGCATCAAG ACCCAGGGTG CCCGTCAGAA GTTCTCCAGC CTCTACATCT
3701 CTCAGTTTAT CATCATGTAT AGTCTTGATG GGAAGAAGTG GCAGACTTAT
3751 CGAGGAAATT CCACTGGAAC CTTAATGGTC TTCTTTGGCA ATGTGGATTC
3801 ATCTGGGATA AAACACAATA TTTTAAACCC TCCAATTATT GCTCGATACA
3851 TCCGTTTGCA CCAACTCAT TATAGCATTC GCAGCACTCT TCGCATGGAG
3901 TTGATGGGCT GTGATT TAAA TAGTTGCAGC ATGCCATTGG GAATGGAGAG
3951 TAAAGCAATA TCAGATGCAC AGATTACTGC TTCATCCTAC TTTACCAATA
4001 TGTTTGCCAC CTGGTCTCCT TCAAAAGCTC GACTTCACCT CCAAGGGAGG
4051 AGTAATGCCT GGAGACCTCA GGTGAATAAT CCAAAAGAGT GGCTGCAAGT
4101 GGAATTCCAG AAGACAATGA AAGTCACAGG AGTAACTACT CAGGGAGTAA
4151 AATCTCTGCT TACCAGCATG TATGTGAAGG AGTTCCTCAT CTCCAGCAGT
4201 CAAGATGGCC ATCAGTGGAC TCTCTTTTTT CAGAATGGCA AAGTAAAGGT
4251 TTTTCAGGGA AATCAAGACT CCTTCACACC TGTGGTGAAC TCTCTAGACC
4301 CACCGTTACT GACTCGCTAC CTTCAATTC ACCCCAGAG TTGGGTGCAC
4351 CAGATTGCCC TGAGGATGGA GGTCTGGGC TGCAGGCAC AGGACCTCTA
4401 C
```

FIG. 14B